

九谷焼に用いられる上絵具の色予測システム構築の研究

Research for Construction of Overglaze Painting Color Prediction System for Kutaniyaki Porcelain

木下 裕太[†]
Yuta Kinoshita上田 芳弘[†]
Yoshihiro Ueda坂本 一磨[†]
Kazuma Sakamoto田代 艶和[†]
Etsuto Tashiro佐々木 直哉[‡]
Naoya Sasaki

1. はじめに

九谷焼は、石川県を代表する伝統工芸品の 1 つである。九谷焼業界では、絵付という工程の際に使用される上絵具が、顔料の廃番や材料メーカーの廃業により、これまで使用してきた材料が入手困難な状況である。そのため、代替材料を使用し、これらの上絵具材料と同等のものを確保する必要がある。また、有鉛絵具は有害性の観点から使用を避けて、それと同等な色を再現しなければならない。そこで、本研究では、回帰分析を用いて、材料を混ぜ合わせて完成する色を予測し、最終的にはデータベースにない色を再現するために必要な材料を表示するシステムを開発する。これにより、代替材料が判明するため、有鉛絵具の使用量の減少にも貢献する。

2. 研究概要

本研究では、 $L^*a^*b^*$ 値[1]を色のデータとして取り扱う。 L は明度に関する値で、0 から 100 までの範囲で表され、 L が高いほど白に、低いほど黒に近づく性質を持つ。 a 、 b はそれぞれ彩度に関する値で、それぞれ-128 から+127 までの範囲で表され、 a が高いほど赤に、低いほど緑に近づく、 b が高いほど黄に、低いほど青に近づく性質を持つ。使用するデータは 2 種類あり、1 つは、原材料のデータ、もう 1 つは、原材料のデータを基に、焼成後の測色値 $L^*a^*b^*$ を記録した化学組成のデータである。これらのデータを使用し、モデルの学習を行う。その後、再現性の確認のため、構築したシステムが算出した化学組成のデータを基に、原材料の分量から上絵具を作製して、焼成後に計測する測色値が、再現目標とする色見本の測色値 ($L^*a^*b^*$) と一致するかの比較し、精度を求める。

3. 使用手法

本研究では、主に決定木、ニューラルネットワークの 2 種類の機械学習モデルに加え、3 次元グラフを用いた直線回帰による解析を行う。

3.1 機械学習モデル

学習には決定木モデルと、ニューラルネットワークを用いた。決定木モデルでは、Classification And Regression Trees (CART) アルゴリズムによる解析と、決定木のアンサンブル学習を行うために、gcForest[2]を改良したものを用いた。

3.1.1 CART

CART とは、データセットを特徴量に基づいて二分木の形で分割する決定木のアルゴリズムの 1 つで、各リー

フノードが最終的な予測を算出する。分割の基準として、分類問題では、ジニ不純度 (Gini impurity) やエントロピー (entropy) が、回帰問題では平均二乗誤差 (MSE) 等が用いられる。

3.1.2 gcForest

gcForest とは、multi-grained scanning 層とカスケードフォレストで構成される決定木ベースのアルゴリズムである。カスケード層では、2 つの完全ランダム決定木フォレスト (complete-random tree forest) と、2 つのランダムフォレストを使用する。各層で入力特徴を学習し、前の層の出力を次の層へと渡すアンサンブル学習の 1 つで、ニューラルネットワークを使用せずに、深層学習と似た高次の学習が可能となる。本手法は、各層で新たな特徴を生成し、全ての特徴を次の層に渡すことで、層が深くなるほど高次の特徴を学習することができる。また、本手法の特徴として、大規模な訓練データを必要とするディープニューラルネットワークとは対照的に gcForest は、小規模データの場合でも精度高く動作することや、ハイパーパラメータの設定がディープニューラルネットワークよりも大幅に少ないため、多大な労力が不要なことが挙げられる。本手法は、画像分類、顔認識、感情分析等の分野に用いられている。本研究では、プログラムを回帰分析へと変更したものを用いる。

3.1.3 ニューラルネットワーク

学習する際の目的変数、説明変数の条件は、決定木モデルと同様とし、本手法では、中間層の数と深さ、学習回数、学習率を設定する必要がある。

3.2 直線回帰による解析

機械学習モデルと別のアプローチとして、直線回帰を用いた。 $L^*a^*b^*$ 値を 3 次元空間として、再現する色に必要な上絵具を予測する。手法としては、 $L^*a^*b^*$ 値と原材料のデータを使用し、データ中の $L^*a^*b^*$ 値を後述する図 3 のように 3 次元空間にプロットする。そして、同じ上絵具の組み合わせで 1 つの着色剤のみ異なるデータを結び、直線を引く。次に、データ間で引いたこの直線と、求めたい点とが最短距離となる直線上の点を計算する。そのため、求めたい測色値を入力し、求めた直線上の点の上絵具の割合を入力値に対応する上絵具の組み合わせと配合とする。ただし、データが無い場所の直線は引けないため、その付近の入力値の予測精度は低くなることが予想される。

4. 実験方法

機械学習モデルでは、色を予測するため、上絵具の

[†] 公立小松大学 Komatsu University

[‡] 石川工業試験場 Industrial Research Institute of Ishikawa

データベースを基に、目的変数を色の指標である $L^*a^*b^*$ 値(3種類), 説明変数を上絵具の材料を構成する 44 種類の化学組成の分量となるよう整理し, 訓練データとして扱う. これにより, 説明変数ごとのモデルの予測値を得ることができ, 作成したモデルと真値との誤差を比較する. 本研究での誤差は, 本モデルの実際の運用を想定した範囲とし[3]を参考に, 3種の目的変数を二乗した総和 ΔE が 2.25 以内に収まることを最低条件とする. こうして出力された化学組成の組み合わせを基に原材料の組み合わせを算出する. なお, データの総数は 220 前後なため, 解析には Leave-One-Out クロスバリデーションを用いる.

5. 結果と考察

機械学習モデルと, 位置ベクトルによる解析の結果について記述する. また, 機械学習モデルで学習した際に得た, 説明変数ごとの特徴量重要度の一部を表 2 に示す.

5.1 機械学習モデル

表 1 より, CART, gcForest では, ΔE 以前に $L^*a^*b^*$ それぞれの誤差が条件を満たしていないが, ニューラルネットワークでは, ΔE が 131.63 となった. CART と gcForest では, 条件を満たさなかったデータの数が 5 割を超えていた. ただ, ニューラルネットワークでは, ΔE が条件を満たさなかったデータは, 全体の 1 割未満に収まった. そこで, 使用したデータを参照すると, ΔE を大きくする原因のデータは, 着色剤を 2 種類以上使用しているものと, 白軸という上絵具を含んでいるものだと読み取れた. 後者は, 例外のデータが多いが, 前者は該当するデータも少なく, 例外がないため, その部分のデータを削除し, 学習を再度行うと精度が 12.59 へと改善された. 白軸を含むデータが上手く学習困難な理由として, 他のデータと比べて色の組み合わせの種類が多いため, 上手く特徴を掴み切れなかったためだと考えている. また, 表 2 はモデルの学習時に強く影響した化学組成を与えた影響が大きい順に 5 つ挙げたものであり, Co_2O_3 (酸化コバルト(III)), AuO_2 (二酸化金) はどのモデルでも影響を与えていることがわかる.

5.2 直線回帰による解析結果

解析の結果を図 3 に示す. 完成した直線回帰による解析システムに, 例として (50, -20, 0) を入力すると, この色を再現するには青, 白軸の上絵具がそれぞれ 80%, 20%, 着色剤 LES11 がこれら 2 種の上絵具全体に対して 2% 必要であることがわかり, 測色値は (50.3, -21.8, -0.7) となるものと推定された. ここで, $L^*a^*b^*$ ごとの色差を求めると (0.3, -1.8, -0.7) となり, ΔE は 3.8 となる. ここで, データが密集する領域とデータが不足している領域に分けて, ΔE を複数算出し, その平均を取ると, 密集領域では 8.8, 不足領域では 107.4 となった. これは, 前述の CART や gcForest と比較すると良好な結果が得られたといえる. 領域ごとの ΔE を比較すると, データが不足する領域では, 予測精度が低くなるので, データの増やすことが必要であり, 現段階では $L > 80$ の領域が不足領域の例として挙げられる. なお, データが密集する領域でも目標である 2.25 を満たしていないため, そちらのデータも増量することで, 精度の改善が見込める.

表 1 機械学習モデルの解析結果

学習方法	ΔL	Δa	Δb	ΔE
CART	89.4	191.2	164.5	71610.1
gcForest	53.4	83.6	118.9	23977.7
ニューラルネットワーク	4.1	4.6	9.7	31.6

表 2 機械学習モデルの解析結果

重要度	学習方法		
	CART	gcForest	ニューラルネットワーク
1	Co_2O_3	Co_2O_3	AuO_2
2	AuO_2	AuO_2	Co_2O_3
3	HfO_2	SO_3	SiO_2
4	TiO_2	SrO	Fe_2O_3
5	CuO	Rb_2O	SrO

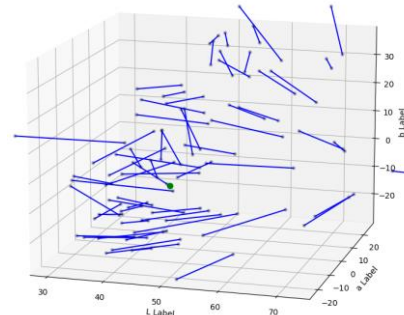


図 3 位置ベクトルによる解析

6. おわりに

本研究では, 様々な手法で材料の組み合わせによる色変化の解析を行った. その結果, ニューラルネットワークを用いた解析が最も精度が高い結果となった. しかし, 実用化にはまだ精度が不十分であるため, 今後は, 精度の向上と同時に, 実際に実証実験を実施して, 予測したモデルの正確性と再現性の確認が必要である. また, データベース以外の組み合わせや, 本研究と異なる実験条件にも対応できるシステムの構築が必要となる.

謝辞

本研究の遂行にあたり, 石川工業試験場の各氏にご意見を賜り, 多大なるご協力を賜りました. ここに記して深く感謝の意を表する.

参考文献

- [1] 株式会社コニカミノルタ. 色々雑学, < <https://www.konicaminolta.jp/instruments/knowledge/color/section2/02.html> > (参照 2023-10-20)
- [2] 深い森(multi-Grained Cascade Forest)と浅い森(Random Forests)を動かして比較してみた.Qiita, < <https://qiita.com/KROYO/items/651bae27feae3e938cf2> > (参照 2024-6-1)
- [3] シーシーエス株式会社. 第 32 回 色差と色均等空間, < https://www.ccs-inc.co.jp/guide/column/light_color/vol32.html > (参照 2023-12-18)