

# メタデータを利用した撮影技術の分類と学習支援 Using Metadata to Classify and Teach Camera Techniques

高橋 一樹

Kazuki Takahashi

法政大学情報科学部デジタルメディア学科

E-mail: kazuki.takahashi.3y@stu.hosei.ac.jp

## Abstract

*In recent years, portable devices such as smartphones are equipped with cameras, and people can easily take photographs. In addition, with the development of SNS, photographs are becoming seen more often. More people want to take better photographs, but it is difficult for them to appropriately set up cameras if they do not have such knowledge. In this research, we develop a system that uses metadata to teach camera techniques by the classification and explanation of how to take good photographs. Specifically, the system first obtains high-quality photographs from the image sharing site Flickr, and classifies camera techniques into six groups by Exif metadata which include the camera setting information at the time of shooting. Next, the system explains camera techniques and necessary equipment based on the classification. We show the result of an experiment on the precision of the following three classification methods: a knowledge-based, a simple perceptron-based, and a multilayer perceptron-based method. Although the precision of these methods is not sufficient, the knowledge-based method is the best among them. Also, we propose the following three methods for supporting users in learning camera techniques by using the classification results: a camera setting screen-based, a camera technique explanation-based, and a quiz-based method. Then we present the result of an experiment of the evaluation of these methods, and show their effectiveness.*

## 1. はじめに

スマートフォンなどの携帯端末にもカメラが搭載されるようになり、写真を撮影することへの敷居が低くなって久しい。また、SNSの発達により、撮影した写真を人に見られる機会、見る機会が増加した。こうした背景から、より上手な写真を撮ることへの欲求が高まっていると考えられる。しかし、今までカメラ任せで撮影してきたユーザーにとって目的に沿ったカメラ設定を行うのは難しい。

本研究では、スマートフォンでマニュアル操作を利用してみたいが、どうしたら良いか解らないユーザーや、画質を求めて一眼カメラを買ったのは良いが、結局オー

トでしか撮影していないようなユーザーが自分自身で適した設定を行えるようになることを目的とし、SNSで人気の写真の撮影方法や必要機材などをメタデータから分類、解説することで撮影を支援するシステムを提案する。具体的には、写真共有 SNS サイト Flickr から質の高い写真を取得し、撮影時のカメラ設定を含んだメタデータである Exif から、注意して撮影を行った点や、使われた撮影技術を推測し、6つのグループに分類する。その後、各分類に基づいた必要な機材や撮影方法を解説することで、ユーザーによる撮影技術の学習を支援する。分類には、知識に基づく方法に加えて、単純パーセプトロンと多層パーセプトロンによる方法も採用し、精度実験を行った。結果、どの分類手法においても精度が十分でなかったものの、知識に基づく方法が最も優れているという結論が得られた。また、分類結果を利用し、カメラ設定方式、撮影技術別表示方式、クイズ方式の3つの撮影技術に関する学習支援方式を提案し、評価実験を行い、有効性を確かめた。

## 2. 関連研究

画像に付与されたメタデータの活用例として、帆足らの研究 [1]が挙げられる。写真に付与された位置情報と Yahoo!ローカルサーチ API、画像検索、類似度算出技術を用いることで、ウェブ上にある情報だけでその画像群が何を写したもののなのか推測できることを示した。

メタデータを利用した撮影者の支援例として、濱田の研究 [2]では、大規模な写真データとその位置情報を利用して、撮影者スポットの選定を支援するツールが開発された。また、渡辺らの研究 [3]で開発された名所カムは、Flickr から取得してきた写真の位置データを利用して、ユーザーの現在地付近で良く写真が撮られている位置を探し、撮影スポットの発見支援を行った。また、撮影した日付と位置情報などを組み合わせた写真管理ソフトを開発した Toyama らの研究 [4]のように、撮影者の支援のみならず、Exifは幅広い研究で活用されている。

以上のように、Exifを利用した研究例は多く存在するが、その多くは位置情報や日時を利用したもので、撮影時のカメラ設定情報に着目したものは非常に少ない。特に、Exifを領した撮影技術の学習支援に関する研究は著者の知る限り行われていない。

### 3. Exif

Exif (Exchangeable image file format)は、デジタルカメラで撮影した写真に埋め込まれるメタデータである。撮影時のカメラ設定のほか、撮影日時や画素数、色空間情報や位置情報などが含まれる。本研究では撮影時の設定情報のうち、写真の写りに大きく影響する3つの項目、シャッター速度、絞り、焦点距離に着目する。

## 4. パーセプトロン

### 4.1. 単純パーセプトロン

単純パーセプトロンは、教師データから学習を行う機械学習の手法である。配列を入力として受け取り、1か-1かの数字を1つ返す。脳細胞の働きを計算機上で表現することを目指した計算モデルであるニューラルネットワークに属し、その中でも最も初期に考案されたモデルである。入力されたパラメータを用いた関数の出力値から、階段関数を用いて、入力されたデータを2つに分類する。線形分離不能な問題は学習できないという欠点がある。

### 4.2. 多層パーセプトロン

多層パーセプトロンは複数のニューロンからなるニューラルネットワークの1種で、複数の単純パーセプトロンを繋ぎ合わせ、1つの単純パーセプトロンの出力を別の入力として扱う。ニューロンが使用する関数は単純パーセプトロンとほぼ同じだが、活性化関数が非連続な階段関数から連続なシグモイド関数に変更されている。

## 5. 提案手法

本研究の提案手法は、Exifを用いた撮影技術の分類と、その結果を用いた学習支援の2つからなる。

### 5.1. 撮影技術の分類

撮影技術の学習に際し、本研究ではExifから把握できる以下の6通りの撮影技術に注目し、対象の写真がどこに属するのか分類する。

- 被写界深度が深い
- 被写界深度が浅い
- 高速シャッターを使用している
- 長時間露光を使用している
- 望遠を使用している
- 広角を使用している

写真をこれら6つのグループへ、重複を許して分類するために、以下の3つ手法を用いる。

- 知識による分類
- 単純パーセプトロンによる分類
- 多層パーセプトロンによる分類

知識による分類は、例えばシャッター速度の値が1/250秒よりも小さければ高速シャッターである、というように、一般に知識として普及している内容に基づき、Exifの値が一定値であればグループに割り振るものである。

今回採用した機械学習には教師データが必要となる。Flickrには特定の撮影機材や方法を用いた写真を投稿するグループ機能が存在するため、ここから取得した撮影

傾向が顕著な写真を教師データとする。入力として使用するのは絞り値、焦点距離、シャッター速度の3つである。単純パーセプトロンは2つのグループに分割する手法であるため、グループごとにパーセプトロンを作成し、対象画像がグループに所属するか、しないかを学習させ、疑似的に多クラス分類を行う。多層パーセプトロンも単純パーセプトロンの場合と同様に教師データを用意し、学習させる。

## 5.2. 撮影技術の学習支援

分類した内容を利用し、撮影技術の学習を支援する3つの方法を提案する。

### 5.2.1. カメラ設定画面方式

1つ目の方式は、一眼カメラの設定画面を模したものである(図1)。各項目にマウスオーバーすることで、その項目を変化させると写真にどのような影響を与えるか、簡単な解説を表示する。また、項目をクリックすることで、対応するグループの作例と撮影技術を解説した画面に移動する。設定画面を模しているため、どこをどう変化させるとどうなるのか、直感的で理解しやすいと考えられる。反面、一眼カメラの設定画面を模しているために、スマートフォンユーザーには解り難いという欠点もあると考えられる。

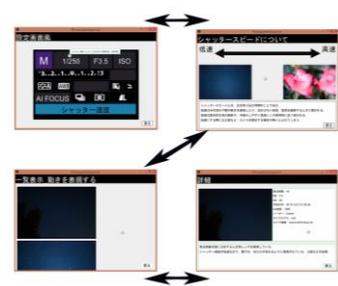


図1 カメラ設定画面方式の画面遷移

### 5.2.2. 撮影技術別表示方式

2つ目の方式は、撮影技術ごとのボタンを作成し、一覧表示するものである(図2)。ボタンにマウスオーバーすることで簡単な解説を表示し、クリックで作例と技術解説画面に移動する。一眼カメラの設定画面に馴染みのないユーザーにも利用しやすく、興味のある撮影技術を探しやすいと考えられる。



図2 撮影技術別表示方式の画面遷移

### 5.2.3. クイズ方式

3 つ目の方式は、撮影意図の推測内容から、重要と思われる Exif の一部を隠し、選択式の問題として出題し、ユーザーに回答させるものである(図 3)。クイズ終了後には正解率や回答率などの結果を表示する。遊びながら学習することができ、自分の実力を試すことができると考えられるが、推測の精度によっては問題として成立しないことやこれ単体で学習するのは難しいことなどが欠点になると考えられる。



図 3 クイズ方式の画面遷移

## 6. 実装

システムの構築には Processing 3.2.3 を用い、Metadata-extractor 2.8.1 等、15 の外部ライブラリを利用した。教師用データとして 6 つの Flickr グループ\* を使用した。それぞれ 500 枚取得し、Exif の欠落したものを除外した最大 3000 枚の写真を用いて学習を行った。実験用のデータとして、分類するグループに対応しそうな Flickr グループ† を選出した。それぞれ 100 枚ずつ取得し、教師用データと同様に Exif の欠落したものを除外した最大 600 枚を対象とし、分類を行った。

## 7. 実験

本節では、実装した分類手法のうち、今回の目的に最も適した手法と、適した手法を用いた学習支援 UI の効果を調査する。

### 7.1. 撮影技術の分類

実装した 3 つの分類手法のうち、最も優れた手法を調査する。分類された内容を写真の知識を持つ被験者 2 名に提示した上で、妥当、妥当でない、解らない、の中から選ばせることで、撮影技術分類の精度を主観的に評価する。

知識に基づいた分類における実験結果を表 1 に示す。全体を通して妥当な結果が得られた。被写界深度が浅いグループの精度が低くなっているのは、望遠や絞りを開いた状態で風景を撮影している写真が含まれたためであ

\* 教師用データは、Hyperfocal DOF technique, Shallow Depth of Field, High Speed Photography 1/10000th Second Flash Exposure Or Faster, LONGEXPOSURE, EF70-200mm f/2.8L IS II USM, Canon EF16-35mm f/2.8L II USM の 6 グループである。

る。また、被写界深度が深いグループにマクロ撮影の写真がわずかに含まれてしまった。どのグループにも属さなかった、無所属なものも全体の 13%程度であり、それをすべて不正解としても 90%程度の精度を誇る。

表 1 知識に基づいた分類の評価

| 被験者A      | 妥当  | 妥当でない | 解らない | 全数  | 精度      |
|-----------|-----|-------|------|-----|---------|
| deep      | 55  | 0     | 1    | 56  | 98.21%  |
| shallow   | 90  | 6     | 9    | 105 | 85.71%  |
| highspeed | 101 | 0     | 0    | 101 | 100.00% |
| slow      | 37  | 1     | 1    | 39  | 94.87%  |
| telephoto | 31  | 0     | 0    | 31  | 100.00% |
| wide      | 93  | 0     | 3    | 96  | 96.88%  |
| 合計        | 407 |       |      | 428 | 95.09%  |
| 無所属       |     |       |      | 34  | 88.10%  |
| 被験者B      | 妥当  | 妥当でない | 解らない | 全数  | 精度      |
| deep      | 58  | 1     | 1    | 60  | 96.67%  |
| shallow   | 108 | 2     | 0    | 110 | 98.18%  |
| highspeed | 98  | 0     | 0    | 98  | 100.00% |
| slow      | 39  | 0     | 2    | 41  | 95.12%  |
| telephoto | 38  | 0     | 0    | 38  | 100.00% |
| wide      | 96  | 0     | 5    | 101 | 95.05%  |
| 合計        | 437 |       |      | 448 | 97.54%  |
| 無所属       |     |       |      | 30  | 91.42%  |

次に、単純パーセプトロンを使用した実験結果を表 2 に示す。分類されているものの多くは妥当だが、どのグループにも所属しないものが多い。また、学習が収束しきっていないからか、同じ対象に実行しても試行するたびに結果が異なり、安定性がない。見当違いなものも含まれていることもある。全体の 35%前後は無所属であり、それを含めると精度は 60%程度になる。クイズの問題数、問題の妥当性に影響が出ると考えられる。

表 5 各学習支援手法の評価

| 被験者A      | 妥当  | 妥当でない | 解らない | 全数  | 精度     |
|-----------|-----|-------|------|-----|--------|
| deep      | 8   | 0     | 1    | 9   | 88.89% |
| shallow   | 51  | 0     | 2    | 53  | 96.23% |
| highspeed | 59  | 17    | 5    | 81  | 72.84% |
| slow      | 35  | 2     | 0    | 37  | 94.59% |
| telephoto | 9   | 0     | 3    | 12  | 75.00% |
| wide      | 4   | 2     | 2    | 8   | 50.00% |
| 合計        | 166 |       |      | 200 | 83.00% |
| 無所属       |     |       |      | 84  | 58.45% |
| 被験者B      | 妥当  | 妥当でない | 解らない | 全数  | 精度     |
| deep      | 5   | 1     | 1    | 7   | 71.43% |
| shallow   | 40  | 2     | 0    | 42  | 95.24% |
| highspeed | 74  | 2     | 0    | 76  | 97.37% |
| slow      | 36  | 4     | 0    | 40  | 90.00% |
| telephoto | 7   | 1     | 2    | 10  | 70.00% |
| wide      | 3   | 0     | 2    | 5   | 60.00% |
| 合計        | 165 |       |      | 180 | 91.67% |
| 無所属       |     |       |      | 92  | 60.66% |

最後に、多層パーセプトロンを用いた分類の結果を表 3 に示す。単純パーセプトロンの結果と同様に所属しているものの精度は良いが、無所属のものがより多い結果となった。こちらも実行するたびに結果が変わってしま

†実験用データは、Japanese Landscape - 日本の風景写真 -, Bokeh and Bokeh 散景世界, High speed, Long Exposure Junkies, Nikon AF-S Nikkor 70-200mm f/2.8G ED VR II, Nikkor 16-35mm f/4 の 6 グループである。

い、時には所属数 0 のグループが出てきてしまう場合もある。

表 3 多層パーセプトロンの評価

| 被験者A      | 妥当 | 妥当でない | 解らない | 全数  | 精度      |
|-----------|----|-------|------|-----|---------|
| deep      | 1  | 0     | 0    | 1   | 100.00% |
| shallow   | 22 | 1     | 0    | 23  | 95.65%  |
| highspeed | 9  | 0     | 0    | 9   | 100.00% |
| slow      | 5  | 1     | 0    | 6   | 83.33%  |
| telephoto | 7  | 1     | 0    | 8   | 87.50%  |
| wide      | 10 | 0     | 1    | 11  | 90.91%  |
| 合計        | 54 |       |      | 58  | 93.10%  |
| 無所属       |    |       |      | 200 | 20.93%  |
| 被験者B      | 妥当 | 妥当でない | 解らない | 全数  | 精度      |
| deep      | 0  | 0     | 0    | 0   | 0.00%   |
| shallow   | 11 | 1     | 0    | 12  | 91.67%  |
| highspeed | 3  | 0     | 0    | 3   | 100.00% |
| slow      | 7  | 1     | 0    | 8   | 87.50%  |
| telephoto | 11 | 3     | 1    | 15  | 73.33%  |
| wide      | 2  | 1     | 2    | 5   | 40.00%  |
| 合計        | 34 |       |      | 43  | 79.07%  |
| 無所属       |    |       |      | 210 | 13.44%  |

以上の結果から、知識に基づいた分類が最も今回の目的に適していると判断した。以降の実験では知識に基づいた分類を利用している。

### 7.2. 撮影技術の学習支援

写真に関する知識のない被験者 6 名を対象に、クイズを 150 問回答させて正解率を記録する。その後、他の学習支援案を受けさせた後、もう一度クイズを 150 問解かせ、正解率の変化で評価を行う。3 名にはすべての学習支援案を、残り 3 名には 3 つの学習手法のうち 1 つのみを用いて学習を行い、各手法の評価を行う。

まず、実装したすべての学習支援手法を使用した場合の正解率の変化を表 4 に示す。学習後には正解率が大きく伸びていることが見て取れる。被験者 a は、今回の被験者の中で最も多く正解を出したが、学習後のクイズで問題の傾向を読まれ、実質 2 択問題となってしまったことが影響している可能性がある。どの被験者も、どの値を極端に設定するとどのように写真に影響するのかが解るようになった。

表 4 学習支援手法全体の評価

| 被験者a | 事前クイズ正解数 | 事前クイズ正解率 | 学習後クイズ正解数 | 学習後クイズ正解率 |
|------|----------|----------|-----------|-----------|
|      | 11       | 7.33%    | 133       | 88.67%    |
| 被験者b | 事前クイズ正解数 | 事前クイズ正解率 | 学習後クイズ正解数 | 学習後クイズ正解率 |
|      | 8        | 5.33%    | 123       | 82.00%    |
| 被験者c | 事前クイズ正解数 | 事前クイズ正解率 | 学習後クイズ正解数 | 学習後クイズ正解率 |
|      | 14       | 9.33%    | 103       | 68.67%    |

次に、撮影技術別表示方式、設定画面風方式、クイズ方式のみを利用して学習した場合の正解率の変化を表 5 に示す。片方の手法だけでも効果が認められる。設定画面風方式を利用した被験者 e は一眼カメラの使用経験はなかったが、こちらのみ実装した項目の説明画面が解りやすいと好評だった。

被験者 f に行ったクイズ方式のみの実験では、事前クイズと学習後クイズの間にクイズを解き続けて貰った。ある程度の効果は認められるが、クイズのみでの学習は難しい。

### 8. 議論

今回採用した分類手法でも精度は 90%程度であり、ク

| 被験者d        | 事前クイズ正解数 | 事前クイズ正解率 | 学習後クイズ正解数 | 学習後クイズ正解率 |
|-------------|----------|----------|-----------|-----------|
| 撮影技術別表示方式のみ | 6        | 4.00%    | 80        | 53.33%    |
| 被験者e        | 事前クイズ正解数 | 事前クイズ正解率 | 学習後クイズ正解数 | 学習後クイズ正解率 |
| 設定画面方式のみ    | 25       | 16.67%   | 94        | 62.67%    |
| 被験者f        | 事前クイズ正解数 | 事前クイズ正解率 | 学習後クイズ正解数 | 学習後クイズ正解率 |
| クイズ方式のみ     | 12       | 8.00%    | 32        | 21.33%    |

イズとして全問解こうと考えると妥当でない問題が登場してしまう。これは以降の実験の正解率にも影響を与えていると考えられる。どの被験者も絞りと被写界深度の概念が理解しにくかったようで、この点の学習支援にも改善の余地がある。本実験で使用した機械学習手法は、精度があまり高くならなかった。バイアスパラメータと学習率の調整や、実装方法の変更、別の手法の利用等によって向上すると考えられる。

### 9. おわりに

本研究では写真の Exif を用いたグループへの分類と、それを利用した撮影技術の学習支援システムを提案した。また、実験よりどの項目がどう写真に影響するかを学習するには有効なシステムであることを示した。今回のシステムでは、絞り値、焦点距離、シャッター速度の 3 つの Exif に特に着目して分類を行ったが、実際にはこれ以外にも写真の写りに大きく影響する要素は多数存在する。被写体との距離による映り方の違いや、センサーサイズによる焦点距離の変化などである。センサーサイズの問題に関しては、カメラを多く扱っているショッピングサイトなどからセンサーの規格を取得し、既定の倍率を掛けることで 35mm 換算数値とすることで解決できる。被写体との距離は現在の Exif から認識することが難しく、Exif に被写体までの距離情報が追加されるのを待つか、画像認識など別のアプローチが求められる。

### 文 献

- [1] 帆足啓一郎, 上向俊晃, 松本一則, 滝嶋康弘, "画像類似度を利用した位置情報付き写真データへのランドマーク情報付与手法," 情報処理学会論文誌, vol. 2, no. 3, pp. 41-52, 2009.
- [2] 濱田康平, "撮影スポットの選定を支援する視覚的ツールの開発(卒業論文)," 筑波大学情報学群情報科学類, 2014.
- [3] 渡辺千穂, 後藤孝行, 塚田浩二, 椎尾一郎, "名所カム : 撮影スポットを探し出すカメラ," WISS 論文集, pp. 168-170, 2010.
- [4] K. Toyama, R. Logan and A. Roseway, "Geographic Location Tags on Digital Images," *Proc. ACM Multimedia*, pp. 156-166, 2003.