

スマートデバイスによる物品探索システムのための 画像弁別性の数値モデル化

Numerical modeling of image discriminability for
home storage and organization system on a smart device

福井 類 前田 啓輔 渡邊 匡彦 下坂 正倫 佐藤 知正
Rui Fukui Keisuke Maeda Masahiko Watanabe Masamichi Shimosaka Tomomasa Sato

東京大学 大学院情報理工学系研究科 知能機械情報学専攻
Mechano-Informatics Course, Graduate School of Information Science and Engineering

Home storage and organization system on smart devices has gained popularity in ubiquitous computing. In our system, thumbnail pictures (Tag Images) of daily-use objects are taken by smart devices, which is used for objects' indexing. Through tag images, a user accesses objects' information. Discriminability of tag images is very important to realize superior usability. For example, an inadequate tag image requires long time to find out an object or may induce miss selection. In this paper, we have constructed a numerical model of tag image's discriminability. The proposed model is based on simple linear regression from popular image features and their statistics. In addition, web-based data input system has also been developed to collect training data efficiently. Consequently, the input system has acquired substantial number of data and a numerical model has been constructed. The constructed model has substantially good performance, however further studies are required to improve it.

1. はじめに

近年、スマートデバイスにおけるタグ画像ベースの所有物管理システム・サービスが複数開発されている [Murakami 08, Torikian 10]. タグ画像ベースシステムとは、目印となる画像(タグ画像)と物品の情報を関連付け、所有物の情報を管理するシステムである(図1). スマートデバイスの普及に伴ってタグ画像の日常的な撮影が可能となり、今後の普及が期待される。しかし、タグ画像の見分けやすさが低いと、一覧表示されたタグ画像の中から目的の画像を探し出すのに時間がかかる、異なるタグ画像と誤認する、といった問題が発生すると考えられる(図2).

そこで本研究では、他の画像と識別し、探し出すことが容易である画像の性質を「画像弁別性」と定義し、画像特徴量にもとづき画像弁別性を数値モデル化することを目指す。これにより、新たな画像の弁別性をユーザに示し、弁別性の高いタグ画像を撮影することを支援するもしくは促す(図3).

本研究では家庭内の物品と収納場所の画像を関連付けて保存し、所有物をどこに収納したのか管理する、スマートデバイスにおける物品探索システムを想定する。そのシステムにおける家庭内の物品および収納場所の画像について画像弁別性の数値モデルを構築する。

2. 数値モデル化の課題と解決策

画像弁別性数値モデル化の課題と解決アプローチを以下に整理する。

1. 画像弁別性と画像特徴量との関係を明らかにすること
⇒ 線形回帰モデルの利用
2. スマートデバイスで計算可能であること
⇒ L1 正則化による画像特徴量の絞り込み
3. 多くの対象に対し、広く再現性を持っていること
⇒ Web ベースシステムによる画像弁別性データの収集

連絡先: 前田 啓輔, 東京大学, 東京都文京区本郷,
maeda@ics.t.u-tokyo.ac.jp



図1: タグ画像ベースシステム 図2: 弁別性の低いタグ画像の例 図3: 弁別性提示による支援の例

2.1 画像弁別性と画像特徴量との関係の明確化

画像の性質を数値化する方法として、画像特徴量という考え方が一般的である。画像弁別性を数値モデル化するためには、画像特徴量との関係を明らかにする必要がある。数値モデル化するための手法として、線形回帰、ロジスティック回帰、ニューラルネットワークなど多くの手法があるが、有用な特徴量の選定が可能であるという利点から本研究では線形回帰モデルを利用する。線形回帰モデルでは式1に表わされるように、画像弁別性を数値化した画像弁別性指標 y を、入力変数である画像特徴量ベクトル \mathbf{x} と重みパラメータ \mathbf{w} の線形和で表現し、訓練データに対して二乗誤差関数 $J(\mathbf{w})$ が最小となるよう重みパラメータを最適化する(式2)。Nは訓練データの数であり、 t_n は教師データである、

$$y(\mathbf{x}, \mathbf{w}) = w_0 + w_1x_1 + \dots + w_Dx_D \quad (1)$$

$$J(\mathbf{w}) = E_D(\mathbf{w}) = \frac{1}{2N} \sum_{n=1}^N \{t_n - \mathbf{w}^T \mathbf{x}_n\}^2 \quad (2)$$

このとき最適化された重みパラメータは各画像特徴量と画像弁別性指標の関係の強さを表し、特徴量の有用性を判断する手が

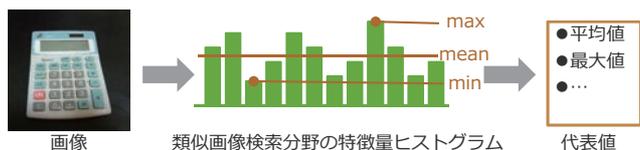


図 4: 特徴量抽出の流れ

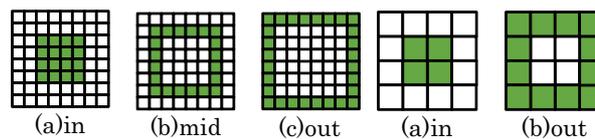


図 5: ColorLayout 領域

図 6: EdgeHistogram 領域

領域

かりとなる。

2.2 スマートデバイスで計算可能な計算

スマートデバイス上のアプリケーションで数値モデルを計算することを考え、計算量を軽量化する必要がある。そこで画像特徴量の次元数を削減する手法を用いる。本研究では数値モデルの入力変数として 1000 次元近い画像特徴量を考慮することになるが、画像特徴量をすべて用いるのではなく、Lasso (least absolute shrinkage and selection operator) [Tibshirani 94] と呼ばれる、特徴量ベクトルの絶対値 $\|\mathbf{w}\|_1$ の和を正則化項として評価関数に加える手法を適用し、重要性の低い画像特徴量を削減する。

2.3 多くの対象に対する広い再現性

画像弁別性の数値モデルは家庭内の多くの物品および収納場所の画像に対してその弁別性を再現できることが求められる。線形回帰モデルを用いて重みパラメータの最適化を行うためには、訓練データとして、画像特徴量と画像弁別性指標を組合わせた画像弁別性データを得る必要がある。線形回帰モデルの訓練データとして用いるため、家庭内の物品と収納場所の多様な画像を収集すると共に、それぞれの画像の弁別性を数値化した指標を算出し、画像弁別性データセットを構築する。本研究では相対的にしか求められない画像弁別性を数値指標として抽出するために一対比較法を用いる。また多数の試行を必要とする一対比較法において、効率的にデータを集めるため、比較対ごとの重要性の考慮と複数人同時入力可能な Web ベースデータ入力システムを構築し利用する。

3. 画像特徴量の設計

本研究では対象画像の特徴を表す指標として、図 4 に示すように 2 段階に分けて抽出する。

まず第 1 段階では、近年目覚ましい発展を遂げている類似画像検索 [Smeulders 00] を参考にした画像特徴量を抽出し、各特徴量ごとにヒストグラムを作成する。具体的には色の配置を表す ColorLayout (MPEG-7 [Chang 01])、エッジの配置を表す EdgeHistogram (MPEG-7 [Chang 01])、エッジの色を表す CEDD [Chatzichristofis 08a]、テキストの色を表す FCTH [Chatzichristofis 08b]、RGB 表現の色分布を表す RGB カラーヒストグラム、HSV 表現の色分布を表す HSV カラーヒストグラムを抽出する。

次に第 2 段階では、第 1 段階で抽出した各特徴量ヒストグラムの代表値を算出して特徴量とするのに加え、画像内の相対位置を表現する必要がある特徴量では画像を複数の領域に分割し、各領域ごとに代表値を算出してその差分も特徴量とする。具体的にはヒストグラムの代表値として平均値、最大値、最小値、範囲、標準偏差、尖度、歪度、最大値位置、最小値位置、最大値と次に大きい位置の差を用い、画像の領域分割としては、ColorLayout を外周、中間、中央の 3 領域、EdgeHistogram を外周、中央の 2 領域に分割し図 5、図 6、それぞれの領域におけるヒストグラムの代表値の差分を特徴量に加える。

4. L1 正則化による画像特徴量の絞り込み

誤差関数 $J(\mathbf{w})$ に正則化項と呼ばれる項を次式のように追加し、モデルの複雑度・自由度に抑制を加え、過学習を防ぐ手法を用いる。 λ はデータに依存する誤差 $E_D(\mathbf{w})$ と正則化項の相対的な重要度を制御する正則化係数である。本研究では Lasso と呼ばれる、式 3 で表されるような、重みパラメータ $\mathbf{w} = \{w_0, \dots, w_D\}$ の絶対値の和を正則化項として用いる。

$$J(\mathbf{w}) = E_D(\mathbf{w}) + \lambda \sum_{i=1}^D |w_i| \quad (3)$$

Lasso は線形回帰の平均二乗誤差最小化を規範とする線形回帰において L1 正則化を加えた学習法であり、正則化項の重みが十分に大きければ、いくつかの特徴量の係数 w_i が 0 になり、疎な解が得られることが知られている。重みパラメータが 0 の特徴量を削減することで、過学習を防ぐとともに計算量の軽量化を図る。

5. 画像弁別性データセットの構築

画像弁別性データセットを構築するために、家庭物品の画像 150 枚と家庭の収納場所の画像 50 枚を用意し、弁別性を学習・評価する対象とした。

画像弁別性を数値化するための手段として、一対比較法を用いた。一対比較法とは、評価したい要素集合の中からランダムに選んだ 2 対象を比較する試行を多数繰り返すことで、要素同士の順位と間隔を評価する手法である。一対比較法は、(1) 順序尺度を間隔尺度に変換できる、(2) 繰り返し行った場合に再現性が高い、(3) 評価方法が容易で負担が少ない、といった利点がある。間隔尺度とは、年齢、温度のように大小関係に加えその差にも意味がある尺度であり、順序尺度とは、順位のように大小関係のみに意味がある尺度である。その一方で、高い精度で評価するためにはすべての組み合わせを複数回ずつ、つまり要素が N あるとき $O(N^2)$ の比較を行うことが推奨されている。しかしながら、すべての画像の組み合わせを同様に比較するのは効率的ではない。

そこで本研究では、各要素の評価値から比較しようとしている組み合わせの勝率を予測し、結果として得られるであろう情報量の多い比較対を優先的に実験協力者に評価してもらうことにより、比較回数を軽減する枠組みを導入した。

要素 i が要素 j よりも優れていると評価される確率が $p_{i,j}$ のとき、比較対 i, j の結果がもつ情報量 $I(i, j)$ は式 4 で表され、評価値が離れている対象の比較よりも、評価値が近い対象の比較の方が得られる情報量が高いことを意味する。ただし、 $E[X]$ は期待値を表し (式 5)、 $h(x)$ は 2 値エントロピーを表す (式 6)。この $I(i, j)$ の値が大きな組み合わせを、優先的に比較する。

$$I(i, j) = E[h(p_{i,j})] \quad (4)$$

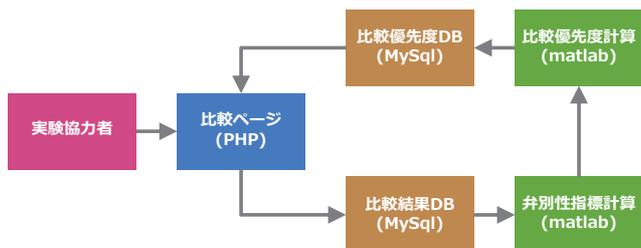


図 7: Web ベースデータ入力システム概要



図 8: 画像比較画面

$$E[X] = \sum_{i=1}^{\infty} x_i P(X = x_i) \quad (5)$$

$$h(x) = x \log \frac{1}{x} + (1-x) \log \frac{1}{1-x} \quad (6)$$

さらに、複数同時試行を行うため、プログラミング言語 PHP によるインタフェースとリレーショナルデータベースライブラリ MySQL によるデータベースをもちいた Web ベースデータ入力システムを構築した。Web ベースデータ入力システムの全体構成を図 7 に示す。大まかな構成は、(1) ユーザインタフェースとなる PHP ページ、(2) 比較優先度を管理するデータベース、(3) 比較結果を保存するデータベース、(4) 過去の比較結果から画像弁別性指標を推定する計算処理、(5) 比較優先度を更新する計算処理、である。図 8 に示すような Web 画面がユーザに示され、どちらかの画像をクリックすると、次の比較対の画像が 2 枚表示されるようになっていく。この時、データ入力者に提示される画像の組み合わせは先の情報量に基づいて自動的に決定されるようになっていく。実験協力者は Web ページ上で優先度の高い比較対を比較し、その結果がデータベースに登録される。これらを用いて複数人が同時にデータ入力をし、結果を集計して比較対の優先度を更新していく流れを図 9 に示す。本研究では 1 日に数回画像の弁別性を算出し、比較優先度データベースを更新した。

実験協力者 62 人により、物品 22000 回、収納場所 3600 回の比較を行った。画像弁別性指標は平均 $\mu = 50$ 標準偏差 $\sigma = 25$ の正規分布に従って分布していると仮定して算出した。

一対比較の結果から画像弁別性指標を推定する手法は Hamza

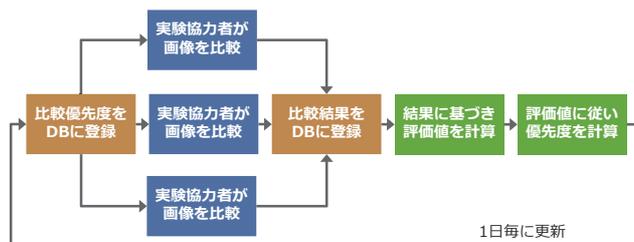


図 9: 複数人同時入力の流れ

	弁別性が高い例	弁別性が低い例
物品		
収納場所		

図 10: 画像弁別性データ例 (右下は評価値)

ら [Aftab 11] を参考にし、各画像弁別性指標の推定値と比較結果それぞれの発生確率が最も高くなるよう、推定値を定めた (式 7)。

$$\hat{\nu} = \{\hat{\nu}_1, \dots, \hat{\nu}_N\} = \arg \max_{\{\nu_k\}} \left(\prod_k P(\nu_k) \right) \left(\prod_{i,j} p_{i,j}^{B(i,j)} \right) \quad (7)$$

$$P(\nu_k) = \frac{1}{\sqrt{2\pi\sigma^2}} \left(\exp -\frac{(\nu_k - \mu)^2}{2\sigma^2} \right) \quad (8)$$

$$p_{i,j} = \frac{1}{2} \left(1 + \operatorname{erf} \left(\frac{\nu_i - \nu_j}{2\sigma} \right) \right) \quad (9)$$

$\hat{\nu} = \{\hat{\nu}_k\} (k = 1, \dots, N)$ は画像弁別性指標の推定値であり、 $P(\nu_k)$ は前述した分布にもとづく画像弁別性指標が ν_k となる確率である (式 8)。また、 $p_{i,j}$ は画像弁別性指標の推定値 ν をもとに算出される、要素 i が要素 j よりも優れていると判断される確率であり (式 9)、 $B(i, j)$ は過去の比較結果において、要素 i が要素 j よりも優れていると判断された回数である。

データ収集の結果、弁別性が高いと評価された画像、弁別性が低いと評価された画像の例を図 10 に示す。画像弁別性が高いと評価された物品画像は明るく鮮明であり、対象物品の輪郭が明確である。逆に画像弁別性が低いと評価された物品画像は、暗く不鮮明であったり、対象物品と背景が判別しにくいことがわかる。画像弁別性が高いと評価された収納場所画像は、明るく鮮明であると同時に目的の収納場所が一意に特定できる。画像弁別性が低いと評価された収納場所画像は、暗く不鮮明であったり、収納場所が一意に特定できないことがわかる。

6. 実験

構築したデータセットに対して重みパラメータを最適化するため、交差検定を行いハイパーパラメータ、Lasso における正則化項の重み λ を決定した。物品画像では $\lambda = 0.44$ 、収納場所では $\lambda = 0.32$ のとき、平均誤差が最小となり、827 次元

表 1: 特徴量の分類

	物品	収納場所
mean	12	14
max	8	13
min	15	7
range	6	6
stdev	20	14
kurtosis	20	11
skewness	22	20
color distance	11	8
max index	7	2
min index	6	4

表 2: 二重交差妥当化による精度評価

	平均絶対誤差	二乗平均平方誤差
物品画像	14.92 ± 0.31	18.81 ± 0.31
収納場所画像	9.72 ± 0.22	16.56 ± 0.13

あった画像特徴量は Lasso により、物品画像で 128 次元、収納場所画像で 99 次元に絞り込まれた。

物品画像において線形回帰モデルの重みが大きくなった特徴量は、画像の中央部と外周部での特徴量の差に関するものが多かった。これは、物品と背景がはっきり分かれている画像の弁別性が高いと評価されたことを意味する。

また、0 でない重みパラメータを有した画像特徴量を設計する際に用いた代表値ごとに分類した概要を表 1 に示す。画像の種類ごとに、上位 3 種を太字にして示した。どの画像でも、標準偏差 (stdev)、歪度 (skewness) が高くなっていることがわかる。これらの統計量はヒストグラムの分布形状を示す特徴量として知られており、標準偏差は分布の散らばり具合、歪度は分布の偏りを示す。すなわち、画像において位置的にあるいは色彩的な分類でヒストグラムを作成した場合に、特徴量の分布が偏っていることが重要であると考えられる。

数値モデルに対して二重に交差検定を行い精度を評価した結果を表 2 に示す。画像弁別性が標準偏差 25 の正規分布に従っていることを考慮すると、平均絶対誤差が物品画像で 14.76、収納場所画像で 9.57 であるので十分な精度で再現できていると考えられる。また、データ収集の結果得られた画像弁別性指標を横軸に、数値モデルから算出した推定値を縦軸にとった相関図を図 11 に示す。左下から右上にかけて分布しており、再現性を持っていることが見て取れる。0 に近い領域では正解よりも大きく、100 に近い領域では正解よりも小さく推定される傾向があることがわかる。これは、パラメータを最適化する上でサンプルが多く分布する平均付近の影響が大きくなったためだと考えられる。

7. おわりに

本研究では画像弁別性数値モデル化の実現を目標に物品と収納場所の画像弁別性データセットを構築し、線形回帰により画像弁別性を定式化し、収集方法も提案した。物品画像において画像の中央部と外周部の特徴量の差が弁別性にとって重要であり、物品、収納場所にかかわらず画像特徴量の分布が偏っていることが弁別性に大きな影響を及ぼすことを示した。データセットに対し画像弁別性を再現できていることを確認した。

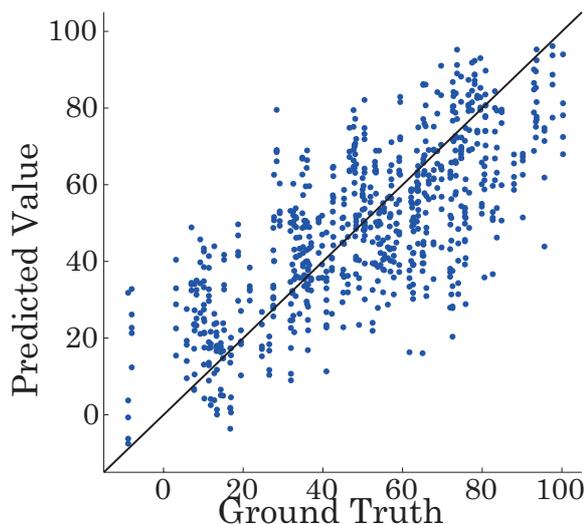


図 11: 推定結果と正解の相関図

参考文献

- [Aftab 11] Aftab, H., et al.: Mutual information scheduling for ranking, in *Proc. FUSION 2011*, pp. 1–8 (2011)
- [Chang 01] Chang, S.-F., et al.: Overview of the MPEG-7 standard, *IEEE Trans. Circuits Syst. Video Technol.*, Vol. 11, No. 6, pp. 688–695 (2001)
- [Chatzichristofis 08a] Chatzichristofis, S. A., et al.: CEDD: color and edge directivity descriptor: a compact descriptor for image indexing and retrieval, in *Proc. ICVS 2008*, pp. 312–322, Springer-Verlag (2008)
- [Chatzichristofis 08b] Chatzichristofis, S. A., et al.: FCTH: Fuzzy Color and Texture Histogram - A Low Level Feature for Accurate Image Retrieval, in *Proc. WIAMIS 2008*, pp. 191–196, IEEE (2008)
- [Murakami 08] Murakami, T.: ItemShelf, <http://itemshelf.com/> (2008)
- [Smeulders 00] Smeulders, A., et al.: Content-based image retrieval at the end of the early years, *Trans. on Pattern Analysis and Machine Intelligence 2000*, Vol. 22, No. 12, pp. 1349–1380 (2000)
- [Tibshirani 94] Tibshirani, R.: Regression Shrinkage and Selection Via the Lasso, *Journal of the Royal Statistical Society, Series B*, Vol. 58, pp. 267–288 (1994)
- [Torikian 10] Torikian, G. J.: Shelves, <https://github.com/gjtorikian/Shelves/> (2010)