

感性への数理科学的なアプローチ

Mathematical Approach to Kansei

栗田多喜夫 小林 匠 大津展之

1. はじめに

画像の内容検索では、その内容を表すキーワード等の情報を画像に付与し、それらを検索キーとして検索する方法が一般的である。しかし、この方法では各画像にキーワードを付与するのに非常に多くの労力が必要であり、画像の内容を表現する適切なキーワードを選択することも難しい。そのため画像自体を検索キーとして、類似した画像を検索する方法（類似画像検索）が盛んに研究されている。

商標や絵画等の特定の対象を集めた画像データベースの検索では、画像の物理的特徴だけでなく利用者の主観的な印象を取り入れた検索も必要となる。そこで、学習用の典型的な画像集合に対する各利用者の印象評価結果から、主観的な印象（印象情報）とそれを生じさせた画像の物理的特徴量（画像情報）との関連性をコンピュータに学習させることが重要となる。筆者らは、利用者の主観的印象を反映した類似画像検索を実現するために判別分析を用いる手法⁽¹⁾、数量化IV類に基づく手法⁽¹⁾、正準相関分析を用いる手法⁽²⁾を提案している。そのほか、田邊らにより多次元尺度構成法を用いる手法も提案されている⁽³⁾。

本稿では、利用者の主観的印象と画像の物理特徴との関連性を学習するための多変量解析手法について概説する。また、その非線形への拡張として、近年急速に利用が拡大しているカーネル多変量解析手法についても紹介する。

2. 感性情報処理のための多変量データ解析手法

利用者の主観的な印象情報を収集するには、幾つかの方法がある。そのため収集方法に応じたデータ解析手法を考える必要がある。ここでは、代表的な手法として、正準相関分析、及び数量化IV類に基づく手法について概説する。

2.1 正準相関分析による主観的印象の学習

学習用の画像集合の各画像に主観的な印象評価値（量的データ）を付与してもらった場合には、正準相関分析を用いて、画像情報と印象情報を直接関連付ける検索空間を構成できる。今、 N 枚の学習用の画像集合があり、各画像から特徴ベクトル $\{\mathbf{x}_i | i=1, \dots, N\}$ が抽出されているとする。また、学習用の N 枚の画像に対して利用者の主観的印象評価が $\{\mathbf{y}_i | i=1, \dots, N\}$ で表現されているとする。このとき、正準相関分析は、画像情報と印象情報の相関が最大となるような線形写像

$$\mathbf{u} = A^T(\mathbf{x} - \bar{\mathbf{x}}_T) \quad (1)$$

$$\mathbf{v} = B^T(\mathbf{y} - \bar{\mathbf{y}}_T) \quad (2)$$

を求める手法である。ここで、 $\bar{\mathbf{y}}_T$ は主観的印象評価ベクトルのサンプル平均である。最適な係数行列は、一般化固有値問題

$$X^T Y (Y^T Y)^{-1} Y^T X A = X^T X A \lambda \quad (3)$$

$$X^T X (X^T X)^{-1} X^T Y B = X^T Y B \lambda \quad (4)$$

の解として求めることができる。ここで、 $X = [(\mathbf{x}_1 - \bar{\mathbf{x}}_T), \dots, (\mathbf{x}_N - \bar{\mathbf{x}}_T)]^T$ 及び、 $Y = [(\mathbf{y}_1 - \bar{\mathbf{y}}_T), \dots, (\mathbf{y}_N - \bar{\mathbf{y}}_T)]^T$ である。式(1)の線形写像を用いることで、任意の画像に対して利用者の印象を反映した空間での表現が得られる。正準相関分析は、画像情報と印象情報を対等に扱う手法であり、画像をキーとして類似した画像を検索できるだけでなく、印象から画像を検索することや未知の画像の印象を推定する等にも利用可能となる⁽²⁾。

栗田多喜夫 正員：フェロー 産業技術総合研究所脳神経情報研究部門

E-mail takio-kurita@aist.go.jp

小林 匠 学生員 産業技術総合研究所情報技術研究部門

大津展之 産業技術総合研究所

Takio KURITA, Fellow (Neuroscience Research Institute, National Institute of Advanced Industrial Science and Technology, Tsukuba-shi, 305-8568 Japan), Takumi KOBAYASHI, Student Member (Information Technology Research Institute, National Institute of Advanced Industrial Science and Technology, Tsukuba-shi, 305-8568 Japan), and Nobuyuki OTSU, Nonmember (National Institute of Advanced Industrial Science and Technology, Tsukuba-shi, 305-8568 Japan).

電子情報通信学会誌 Vol.92 No.11 pp.917-919 2009年11月

©電子情報通信学会 2009

2.2 数量化IV類に基づく手法による主観的印象の学習

利用者の主観評価としてより詳細な情報は、学習用の画像集合内のすべての画像間の主観的な類似性を計測することであろう。二対比較などにより画像間の類似性を人間に評価してもらうことも可能であるが、画像の枚数が増えると多大な労力が必要となる。一方、各画像に主観的な印象評価値を付与してもらえば、それらから画像間の類似度を求めることもできる。ともかく、何らかの方法で画像間の主観的な類似度が求まった場合には、数量化IV類に基づく手法を用いて、印象情報と画像の物理情報の関連性を学習することができる⁽¹⁾。

数量化IV類は、 N 個の対象（ここでは画像）間の類似度 $\{s_{ij} | i, j=1, \dots, N\}$ が与えられたとき、類似度の大きい画像同士は近くに、類似度の小さい画像同士は遠くなるように各画像をユークリッド空間内の多次元ベクトル $\{\mathbf{u}_i | i=1, \dots, N\}$ として配置する手法である。そのため、評価関数として、

$$Q = -\frac{1}{2} \sum_{i \neq j} s_{ij} \|\mathbf{u}_i - \mathbf{u}_j\|^2 \quad (5)$$

を考える。ベクトル \mathbf{u} の大きさに関する制約条件を置くと、最適な多次元ベクトル $U = [\mathbf{u}_1, \dots, \mathbf{u}_N]^T$ は、類似度から定義される行列 $L = D - S$ の固有値問題

$$LU = UA \quad (U^T U = I) \quad (6)$$

の解として求められる。ただし、 $D = \text{diag}(\sum_j s_{ij})$ 、 $S = [s_{ij}]$ である。得られた多次元ベクトル $\{\mathbf{u}_i | i=1, \dots, N\}$ は、利用者の主観的な類似度を反映するように空間配置される。

しかし、この空間では感性を生じさせる対象の物理量との関連性については表現されていない。そこで、各画像から抽出した特徴ベクトル \mathbf{x} と類似度を反映した空間配置との間に線形モデル $\mathbf{u} = A^T(\mathbf{x} - \bar{\mathbf{x}}_T)$ を仮定し、画像の特徴ベクトルから類似度を反映した空間配置を構成することを考えてみよう。この場合には、評価関数(5)を最大とするような最適な係数行列 A は、上記の手法と同様に、一般化固有値問題

$$X^T L X A = X^T X A A \quad (A^T (X^T X) A = I) \quad (7)$$

の解として求めることができる。この線形モデルを用いることで画像の特徴ベクトルと利用者の主観的な類似度を反映する空間とを関連付けることが可能となり、利用者の主観的な類似度を反映した画像検索が実現できる⁽¹⁾。

3. カーネル多変量解析手法

前章で紹介した手法は、どれも画像情報と印象情報の

関係を線形でモデル化する手法である。しかし、画像情報と印象情報の関係は非線形であるとも容易に想像でき、線形モデルで近似することは一般には難しい。最近注目されているカーネル法⁽⁴⁾は、カーネル関数の重み付きの和で表したモデルを正則化付きで最適化するデータ解析手法であり、前章で紹介した手法を非線形に拡張することを可能にする。

3.1 カーネル正準相関分析

正準相関分析はカーネル法を用いて非線形に拡張することができる。式(1)、(2)に対応する写像を画像情報のカーネル関数と印象情報のカーネル関数を用いて、

$$\mathbf{u}(\mathbf{x}) = \sum_{i=1}^N \alpha_i k_X(\mathbf{x}_i, \mathbf{x}) = A^T \mathbf{k}_X(\mathbf{x}) \quad (8)$$

$$\mathbf{v}(\mathbf{y}) = \sum_{i=1}^N \beta_i k_Y(\mathbf{y}_i, \mathbf{y}) = B^T \mathbf{k}_Y(\mathbf{y}) \quad (9)$$

でモデル化する。正則化を行った場合の最適な係数行列は、

$$K_X J K_Y (K_Y J K_Y + \zeta_Y K_Y)^{-1} K_Y J K_X A \\ = (K_X J X_X + \zeta_X K_X) A A \quad (10)$$

$$K_Y J K_X (K_X J K_X + \zeta_X K_X)^{-1} K_X J K_X B \\ = (K_Y J K_Y + \zeta_Y K_Y) B A \quad (11)$$

の解として求めることができる⁽⁵⁾。ここで、 K_X 、 K_Y 、 ζ_X 及び ζ_Y は、それぞれ、画像情報に対するカーネルのグラム行列、印象情報に対するカーネルのグラム行列、画像情報の正則化パラメータ、及び、印象情報の正則化のパラメータであり、 $J = I - \frac{1}{N} \mathbf{1}\mathbf{1}^T$ である。

3.2 カーネル数量化IV類に基づく手法

数量化IV類に基づく手法をカーネル法に基づいて非線形に拡張する。ここでは、カーネル関数を用いた非線形写像 $\mathbf{u}(\mathbf{x}) = A^T \mathbf{k}_X(\mathbf{x})$ により、画像の特徴ベクトルから類似度を反映した空間配置を構成することを考える。最適な係数行列は、やはり、一般化固有値問題

$$K_X L K_X A = K_X K_X A A \quad (12)$$

の解として求めることができる。

数量化IV類は、類似度の定義を変更することで、様々な距離尺度を反映させることができる統一的な手法である。実際、数量化IV類と本質的に同じであるが、サンプル間の近さ（類似度）をガウスカーネル等のカーネル関数を用いて定義し、その近さを反映した空間配置を求める手法は、ラブラシアン固有マップ法と呼ばれている⁽⁵⁾。

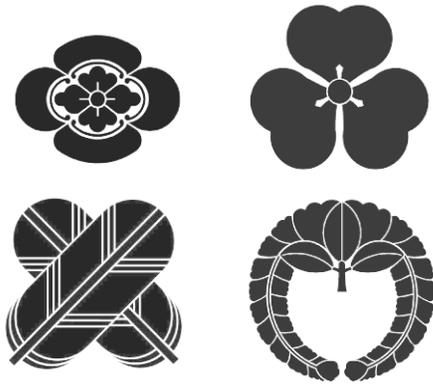


図1 家紋の例

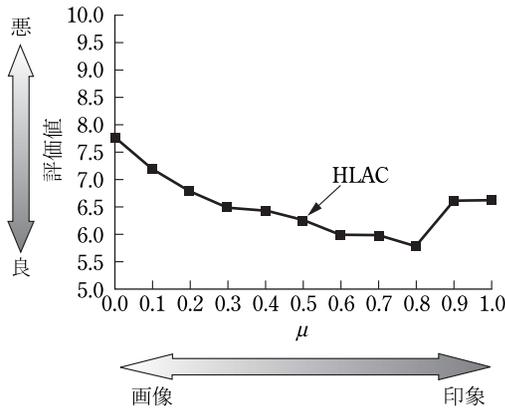


図2 印象割合の比較実験の結果

4. 主観的印象からの類似画像検索

ここでは、家紋を例に、画像情報と印象情報の両方を加味した検索実験について紹介する⁽⁶⁾。図1に家紋の例を示す。画像情報としては、高次局所自己相関特徴(HLAC)⁽⁷⁾を用いた。印象情報としては、24名の被験者に各家紋を見せて、「あでやかな」、「上品な」、「エレガントな」のような26語の印象語に対する印象を5段階の評定尺度によって回答してもらった結果を利用した。

カーネル数量化IV類に基づく手法は、印象情報と画像情報を統合した類似度

$$s_{ij} = \mu \exp\left(-\frac{\|\mathbf{y}_i - \mathbf{y}_j\|^2}{2\sigma_y^2}\right) + (1-\mu) \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma_x^2}\right) \quad (13)$$

を用いることにより、印象情報と画像情報を割合 μ でブレンドしたような検索空間を構成できる。ここで、 $\mu=1$ の場合は、印象情報を最大限反映させた空間を構成することになり、 $\mu=0$ は印象情報を全く用いなくて画像情報のみで検索することになる。

50個の家紋に対して検索性能を leave-one-out 法で評

価した結果を図2に示す。この結果から、最良の評価値は、 $\mu=0.8$ のときに得られている。つまり、印象情報だけでなく、画像情報も加味した類似度に基づき検索空間を構成することで、検索性能が向上することを示している。

5. まとめ

本稿では、主観的印象からの類似画像検索を例に、感性情報と感性を生じさせる対象の物理量との関係を学習する手法について紹介した。カーネル法等の非線形の特徴解析手法が利用できるようになり、静的で再現性のあるような感性情報については、モデル化が可能になりつつある。今後は、状況や意図に応じた感性情報の扱いや動的に変動する感性への追従等の難しい課題にも挑戦すべき時期がきたように思う。

文 献

- (1) 栗田多喜夫, 下垣弘行, 加藤俊一, “主観的類似度に適応した画像検索,” 情処学論, vol.31, no.2, pp. 227-237, 1990.
- (2) 栗田多喜夫, 加藤俊一, 福田郁美, 坂倉あゆみ, “印象語による絵画データベースの検索,” 情処学論, vol.33, no.11, pp.1373-1383, 1992.
- (3) 田邊勝義, 大谷 淳, 石井健一郎, “多次元心理空間を用いる類似画像検索法,” 信学論(D-II), vol.J75-D-II, no.11, pp.1856-1865, Nov. 1992.
- (4) 赤穂昭太郎, カーネル多変量解析, 岩波書店, 2008.
- (5) M. Belkin and P. Niyogi, “Laplacian eigenmaps and spectral techniques for embedding and clustering,” Advances in Neural Processing Systems (NIPS14), pp.585-591, 2002.
- (6) 小林 匠, 森崎功一, 大津展之, “印象情報の付与による類似画像検索性能の評価,” 信学論(D), vol.J91-D, no.4, pp.1025-1032, April 2008.
- (7) N. Otsu and T. Kurita, “A new scheme for practical flexible and intelligent vision systems,” Proc. of IAPR Workshop on Computer Vision, pp.431-435, 1988.

(平成21年5月30日受付 平成21年7月14日最終受付)



栗田 多喜夫 (正員:フェロー)

昭56名工大・工・電子卒。同年電子技術総合研究所。以来、統計的パターン認識とその画像認識への応用に関する研究に従事。博士(工学)。現在、産業技術総合研究所脳神経情報研究部門副研究部門長、筑波大連携大学院教授。日本神経回路学会、人工知能学会、情報処理学会、IEEE CS 各会員。



小林 匠 (学生員)

平15東大・工・機械情報卒。平17同大学院修士課程了。同年(株)東芝入社。現在、産業技術総合研究所研究員。博士(工学)。在学中よりパターン認識・多変量解析の理論と応用の研究に従事。



大津 展之

昭44東大・工・計数卒。昭46同大学院修士課程了。同年電子技術総合研究所入所。以来、パターン認識、画像処理、多変量データ解析、人工知能に関する数理的基礎研究に従事。工博。現在、産業技術総合研究所フェロー、筑波大連携大学院教授。日本行動計量学会、情報処理学会等会員。