

幾何データからの Geometric Algebra を 用いた特徴抽出

Pham Minh Tuan¹ · 橋 完太¹ · Eckhard Hitzer² · Sven Buchholz³ ·
吉川 大弘¹ · 古橋 武¹

(受付 2008 年 1 月 7 日; 改訂 2008 年 4 月 22 日)

要 旨

従来のデータマイニングにおけるパターン認識では、空間的な特徴を含むデータに対しては幾何的な性質を考慮せずに特徴を抽出する。本研究では、複素数や四元数の一般形である Geometric Algebra が持つ、空間的な物体や物体間の関係を簡潔に記述できるという長所を活かして、データの空間的な特徴を抽出する。本稿では、空間ベクトルの系列で与えられるデータからの特徴抽出について、Geometric Algebra を用いる系統的な手法を提案する。そして、データが幾何的な特徴を持つ手書き文字の分類問題について、提案手法により従来は偶然または経験的にしか発見・注目できなかった特徴を抽出でき、分類精度および頑健性が向上することを示す。

キーワード： Geometric Algebra, 特徴抽出, 混合正規分布モデル, 混合エキスパート, パターン認識, 手書き数字分類。

1. はじめに

様々な分野で膨大なデータが蓄積されている現在では、データを解析し有力な情報を発掘するデータマイニングがますます必要となっている。しかし、従来のデータマイニングにおけるほとんどのパターン認識手法では、空間的な特徴を含むデータ(以下、幾何データ)に対しても幾何的な性質を考慮せずに特徴抽出する。例えば、ある 3 次元物体に関して m 本の 3 次元空間ベクトルが計測された場合、多くの従来手法では、各ベクトルの 3 座標を順番に m 組ならば、 $3m$ 個の実数、あるいは $3m$ 次元実ベクトル $x \in \mathbb{R}^{3m}$ として扱う。ところが、これらの座標値は座標系の定義に依存するため、たとえば、学習データの計測環境と異なる環境で計測されたデータに関して著しく推論が悪くなるといった問題が生じる。さらに、各変数に正規化などを行った場合には 3 次元空間での形状に関する情報は失われてしまう。これに対して従来手法でも座標系に依存しない特徴抽出が考えられるだろうが、そのような特徴抽出が発案され採用されるか否かは偶然やモデル構築者の経験に依存する。

本研究では、空間ベクトルや空間ベクトル間の関係を簡潔に記述できる Geometric Algebra

¹ 名古屋大学大学院 工学研究科：〒 464-8601 名古屋市中千種区不老町

² 福井大学 物理工学科：〒 910-8507 福井市文京 3-9-1

³ Institut für Informatik, Christian-Albrechts-Universität zu Kiel, Olshausenstrasse 40, D-24098 Kiel, Germany

(GA) (Hestenes, 1999; Doran and Lasenby, 2003; Dorst et al., 2007) を用いて, 幾何データからの系統的な特徴抽出を行い, 分類問題における精度および頑健性の向上を狙う. GA の一部ともみなすことができる複素数や四元数については, これらをそれぞれ用いた画像処理・時系列信号処理や色付画像処理・多次元時系列信号処理などの, 空間的な処理を要する応用での成功例が多く挙げられる. (Sekita et al., 1992; Hirose, 2003; Matsui et al., 2004; Buchholz and Bihan, 2006; Nitta, 1997; Hildenbrand and Hitzer, 2008; Hitzer, 2007) また, GA 値入出力関数の学習手法 (Sommer, 2001) も提案され, その有効性が示されている.

提案手法で GA によって抽出した特徴は, 入出力関数の学習にもデータ分布の学習にも, 用いることができる. 本研究では, 幾何データから GA を用いて系統的に複数抽出する特徴データの分布に対して, 正規分布や混合正規分布モデルを学習させ, 多クラス分類問題へ適用する. 提案手法で導出する複数の特徴抽出には, 座標系に依存するものとしなないものがあるため, 複数 GMM の混合エキスパートの分類問題への適用も提案する. 幾何データの多クラス分類問題の例として, UCI レポジトリ (Asuncion and Newman, 2007) の手書き文字 (数字) データを用いる. 学習済みのモデルを用いて実際の手書き文字 (数字) を分類する際, 分類したいデータの座標系が学習データの座標系とぴたりと一致することは考えにくい. そのため本稿では, テストデータにランダムな回転を加えた場合の分類性能により頑健性を評価する.

2. 手法

2.1 Geometric Algebra

Geometric Algebra (GA) は代数構造の一種であり, クリフォード代数とも呼ばれる. 実ベクトル空間 \mathbb{R}^n には, 正規直交基底 $\mathcal{O} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ を定めることができる. 次数 n の GA は, 本稿では \mathcal{G}_n で表し, \mathcal{O} 相互の結合的で双線形な積演算: 幾何積 (geometric product)

$$(2.1) \quad \mathbf{e}_i \mathbf{e}_j = \begin{cases} 1 & (i=j) \\ -\mathbf{e}_j \mathbf{e}_i & (i \neq j) \end{cases}$$

により構成される. なお, n 個のうち一部または全部の基底が $\mathbf{e}_i^2 = -1$ となる場合も, たかだか各項の符号が変わるだけであり, 以下の議論を同様に展開できる. 本稿では (2.1) の場合のみを考える.

いま, \mathbb{R}^n において 2 本のベクトル $\{\mathbf{a}_l = \sum_i a_{li} \mathbf{e}_i, l=1, 2\}$ を考える. これらの幾何積は

$$(2.2) \quad \mathbf{a}_1 \mathbf{a}_2 = \sum_{i=1}^n a_{1i} a_{2i} + \sum_{i=1}^{n-1} \sum_{j=i+1}^n (a_{1i} a_{2j} - a_{1j} a_{2i}) \mathbf{e}_{ij}$$

となる. ここで $\mathbf{e}_{ij} = \mathbf{e}_i \mathbf{e}_j$ であり, 以下同様に略記する. 第 1 項はベクトル代数と同様の内積 $\mathbf{a}_1 \cdot \mathbf{a}_2$ であり, 内積は対称 $\mathbf{a}_1 \cdot \mathbf{a}_2 = \mathbf{a}_2 \cdot \mathbf{a}_1$ である. 第 2 項は外積 $\mathbf{a}_1 \wedge \mathbf{a}_2$ と呼ばれる. 外積は Grassmann の反対称積とも呼ばれ $\mathbf{a}_1 \wedge \mathbf{a}_2 = -\mathbf{a}_2 \wedge \mathbf{a}_1$ となる. 第 2 項は, 2-ブレードと呼ばれる. 2-ブレードの線形和は, バイベクトルと呼ばれ, $\{1, \dots, n\}$ から 2 つを選ぶ集合を $\mathcal{I}_2 = \{i_1 i_2 \mid 1 \leq i_1 < i_2 \leq n\}$ とすると, $\sum_{I \in \mathcal{I}_2} w_I \mathbf{e}_I, w_I \in \mathbb{R}$ と表される. なお, $\mathbf{a}_1 = \kappa \mathbf{a}_2, \kappa \in \mathbb{R}$ であるとき第 2 項は 0 となる.

次に, $\mathbf{a}_1 \mathbf{a}_2 = \mathbf{a}_1 \cdot \mathbf{a}_2 + \mathbf{a}_1 \wedge \mathbf{a}_2$ と第 3 のベクトル \mathbf{a}_3 との幾何積を考える. まず, $\mathbf{a}_1 \cdot \mathbf{a}_2 \in \mathbb{R}$ であるので, $(\mathbf{a}_1 \cdot \mathbf{a}_2) \mathbf{a}_3$ はベクトルである. 一方, $(\mathbf{a}_1 \wedge \mathbf{a}_2) \mathbf{a}_3 = (\sum_{I \in \mathcal{I}_2} w_I \mathbf{e}_I) \sum_i a_{3i} \mathbf{e}_i$ であって, ある $I = i_1 i_2$ に対して

$$(2.3) \quad \mathbf{e}_I \mathbf{e}_i = \mathbf{e}_{i_1} \mathbf{e}_{i_2} \mathbf{e}_i = \begin{cases} \mathbf{e}_{i_1} & (i = i_2) \\ -\mathbf{e}_{i_2} & (i = i_1) \\ \mathbf{e}_{i_1 i_2 i} & (i_1 \neq i \neq i_2) \end{cases}$$

である。よって $(\mathbf{a}_1 \wedge \mathbf{a}_2) \mathbf{a}_3$ はベクトル(式(2.3)の上2行に該当する部分の総和)と3-ブレード $\mathbf{a}_1 \wedge \mathbf{a}_2 \wedge \mathbf{a}_3$ (式(2.3)の下1行に該当する部分の総和)とに分けられる。3-ブレードの線形和はトライベクトルと呼ばれ、 $\{1, \dots, n\}$ から3つを選ぶ集合 $\mathcal{I}_3 = \{i_1 i_2 i_3 \mid 1 \leq i_1 < i_2 < i_3 \leq n\}$ を用いて、 $\sum_{I \in \mathcal{I}_3} w_I \mathbf{e}_I$ と表される。

以下同様に、互いに独立な $\mathbf{a}_1, \dots, \mathbf{a}_k, (k \leq n)$ の幾何積 $\mathbf{a}_1 \cdots \mathbf{a}_k$ を考える際に、 k -ブレード $\mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_k$ が出現する。任意の k -ブレードの線形和である k -ベクトルは、 $\{1, \dots, n\}$ から k 個を選ぶ集合 $\mathcal{I}_k = \{i_1 \cdots i_k \mid 1 \leq i_1 < \dots < i_k \leq n\}$ を用いて、 $\sum_{I \in \mathcal{I}_k} w_I \mathbf{e}_I$ と表される。また、 $\mathcal{I}_1 = \{i_1 \mid 1 \leq i_1 \leq n\}$ を用いて1-ブレードであるベクトルを、 $\mathcal{I}_0 = \{\emptyset\}$ を用いて0-ブレードであるスカラーを、それぞれ表すことができる。スカラーについては $\mathbf{e}_\emptyset = 1$ と表す。 \mathcal{G}_n の k -ブレード全体を $\wedge^k \mathbb{R}^n$, k -ベクトル全体を \mathcal{G}_n^k と表す。

ここで、 $n \leq 3$ では $0 \leq k \leq n$ についてすべての k -ベクトルは同時に k -ブレードでもある。つまり k 本のベクトルの外積で表すことができる。しかし、 $n \geq 4$ では、 k -ブレードでない k -ベクトルが存在する。例えば $\mathbf{e}_{12} + \mathbf{e}_{34}$ はバイベクトルであるが2本のベクトルの外積として表せないので、2-ブレードではない。 k -ブレードと k -ベクトルの関係を改めて強調すると $\wedge^k \mathbb{R}^n = \{A \in \mathcal{G}_n^k \mid \exists \{\mathbf{b}_1, \dots, \mathbf{b}_k\}, A = \mathbf{b}_1 \wedge \dots \wedge \mathbf{b}_k\}$ である。 \mathcal{G}_n の最も一般的な元 A は、 $\mathcal{I} = \bigcup_{k=0}^n \mathcal{I}_k = \mathcal{P}(\{1, \dots, n\})$ を用いて、 $A = \sum_{I \in \mathcal{I}} w_I \mathbf{e}_I$ と表される。ここで $\mathcal{P}(\cdot)$ はベキ集合を表す。より具体的に、 \mathcal{G}_2 の一般の元は $\{1, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_{12}\}$ の線形和であり、 \mathcal{G}_3 の一般の元は $\{1, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_{12}, \mathbf{e}_{13}, \mathbf{e}_{23}, \mathbf{e}_{123}\}$ の線形和である。また $\langle A \rangle_k = \sum_{I \in \mathcal{I}_k} w_I \mathbf{e}_I$ を用いて、 $A = \sum_{k=0}^n \langle A \rangle_k$ と表される。 $\langle \cdot \rangle_k$ は k -ベクトル部分のみを取り出す演算とみなせる。スカラー部分を取り出す演算は $\langle \cdot \rangle = \langle \cdot \rangle_0$ と略記する。

$k (1 \leq k \leq n)$ 個のベクトルの幾何積をまとめると、

$$(2.4) \quad \begin{aligned} \mathbf{a}_1 &\in \mathcal{G}_n^1 = \mathbb{R}^n = \bigwedge^1 \mathbb{R}^n \\ \mathbf{a}_1 \mathbf{a}_2 &\in \mathcal{G}_n^0 \oplus \bigwedge^2 \mathbb{R}^n \\ \mathbf{a}_1 \mathbf{a}_2 \mathbf{a}_3 &\in \mathcal{G}_n^1 \oplus \bigwedge^3 \mathbb{R}^n \\ &\vdots \\ \mathbf{a}_1 \mathbf{a}_2 \cdots \mathbf{a}_k &\in \begin{cases} \mathcal{G}_n^1 \oplus \mathcal{G}_n^3 \oplus \cdots \oplus \mathcal{G}_n^{k-2} \oplus \bigwedge^k \mathbb{R}^n & (\text{odd } k) \\ \mathcal{G}_n^0 \oplus \mathcal{G}_n^2 \oplus \cdots \oplus \mathcal{G}_n^{k-2} \oplus \bigwedge^k \mathbb{R}^n & (\text{even } k) \end{cases} \end{aligned}$$

となる。ここで集合 \mathcal{A}, \mathcal{B} について $\mathcal{A} \oplus \mathcal{B} \equiv \{a + b \mid a \in \mathcal{A}, b \in \mathcal{B}\}$ である。

2.2 特徴抽出

n 次元での m 本のベクトルの集合 $\xi = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_m\}$, $\mathbf{p}_i = [x_{i,1}, \dots, x_{i,n}]$ から作ることができる特徴の多くは、

$$(2.5) \quad \begin{aligned} F_1^m &= \{x_{1,1}, \dots, x_{1,n}, \dots, x_{m,1}, \dots, x_{m,n}\} \\ F_2^m &= \{x_{1,1}^2, x_{1,1}x_{1,2}, \dots, x_{1,1}x_{m,n-1}, x_{1,1}x_{m,n}, \\ &\quad x_{1,2}^2, \dots, x_{1,2}x_{m,n}, \dots, x_{m,n-1}^2, x_{m,n-1}x_{m,n}, x_{m,n}^2\} \end{aligned}$$

$$F_3^m = \{x_{1,1}^3, x_{1,1}^2 x_{1,2}, \dots, x_{1,1} x_{m,n-1} x_{m,n}, x_{1,1} x_{m,n}^2, \\ x_{1,2}^3, \dots, x_{1,2} x_{m,n}^2, \dots, x_{m,n-1}^3, x_{m,n-1}^2 x_{m,n}, x_{m,n-1} x_{m,n}^2, x_{m,n}^3\} \\ \vdots$$

の線形和 $\mathcal{F}_k = \{\sum_{\phi \in F_k^m} w_\phi \phi \mid w_\phi \in \mathbb{R}\}$ によって表すことができる。しかし特徴の数は

$$(2.6) \quad \begin{aligned} |F_1^m| &= nm = O(nm) \\ |F_2^m| &= \sum_{i=1}^m |F_1^i| = n \frac{m(m+1)}{2} = O(nm^2) \\ |F_3^m| &= \sum_{i=1}^m |F_2^i| = O(nm^3) \\ &\vdots \\ |F_k^m| &= \sum_{i=1}^m |F_{k-1}^i| = O(nm^k) \\ &\vdots \end{aligned}$$

と膨大になり、識別問題においてそれらのすべてが識別に有効な特徴になるとは限らないし、これらのすべてを学習モデルに入力することは計算量の面から現実的でない。そのため、この膨大な特徴空間のうち適切な部分空間を設定する必要がある。従来この設定はアドホックに行われてきた。

従来、識別問題に関する特徴抽出について様々な手法が提案された。例えば、オンライン手書き文字認識での特徴抽出は最も簡単な座標値特徴 (Bahlmann and Burkhardt, 2004; Guyon et al., 1991; Hu et al., 1999) が多く使われる。これは上記の部分特徴空間 \mathcal{F}_1 に他ならない。 \mathcal{F}_1 は計測座標系の回転の影響を受ける。その他 Connell and Jain (2002) はオンライン手書き文字の識別問題に関して、6種類の特徴抽出を提案した。6種類のうち4種類は \mathcal{F}_1 に含まれるが、残りの回転に不変な2種類の特徴のそれぞれは2本のベクトル $\mathbf{p}_{i+1} - \mathbf{p}_i$ と $\mathbf{p}_i - \mathbf{p}_{i-1}$ 間の角度 θ に対する $\cos \theta$ および $\sin \theta$ と本質的に変わらず、 $\cos \theta$ も $\sin \theta$ も \mathcal{F}_2 に含まれる (ペン先の軌道上で \mathbf{p}_i と \mathbf{p}_{i+1} の間隔を変えた2種類を特徴として抽出する)。また、画像認識で用いられるモーメント特徴 (Hu, 1962) および複素モーメント特徴 (Abo-Zaid et al, 1988) は、全てのベクトルに対する一般化慣性モーメントであり、次数 (p, q) のモーメントは $k = p + q$ として \mathcal{F}_k のうち $\{x_{i,1}^p x_{i,2}^q \mid i = 1, \dots, m\}$ の線形和で表される。しかし、これらのモーメントが画像に対して提案された経緯から、モーメントではベクトル間 (画像の場合はピクセル間) の関係は考慮されない。つまり \mathcal{F}_k において $i \neq j$ となる $x_{i,1}^p x_{j,2}^q$ の係数は常に0となる。そこで、本研究では、回転に不変であり、ベクトル間 (画像の場合はピクセル間) の関係を考慮するような特徴も含めた部分空間を GA の幾何積により系統的に抽出する手法を提案する。

空間ベクトルの集合、または、系列 (順序つき集合) $\xi = \{\mathbf{p}_l \in \mathbb{R}^n, l = 1, \dots, m\}$ からの系統的な特徴抽出を以下のように提案する。まず、 ξ が空間ベクトルの (順序なし) 集合である場合を考える。各 k -ブレードと式 (2.4) で $k=2$ のときに出現するスカラーをそれぞれ特徴として抽出するという方針で、空間ベクトルの集合 ξ から、 $n' = \min\{n, m\}$ として、以下の $n' + 1$ 通りの特徴抽出を導出する。 $k = 1, \dots, n'$ について、

$$(2.7) \quad \begin{aligned} f_0(\xi) &= \{\langle \mathbf{p}_{l_1} \mathbf{p}_{l_2} \rangle\} \in \mathbb{R}^{\binom{m}{2} + m} \\ f_k(\xi) &= \{\langle \mathbf{p}_{l_1} \cdots \mathbf{p}_{l_k} \mathbf{e}_I^{-1} \rangle, I \in \mathcal{I}_k\} \in \mathbb{R}^{\binom{m}{k} |\mathcal{I}_k|} \end{aligned}$$

ここで、 \mathbf{e}_I^{-1} は \mathbf{e}_I の逆数であり、 $I = i_1 i_2 \cdots i_k$ の場合 $\mathbf{e}_I^{-1} = \mathbf{e}_{i_k} \cdots \mathbf{e}_{i_2} \mathbf{e}_{i_1}$ である。 $\mathbf{p}_I \mathbf{p}_{I+1} \cdots \mathbf{p}_{I+k-1}$ の \mathbf{e}_I の係数は \mathbf{e}_I^{-1} との幾何積を施した後のスカラー部分として、取り出すことができる。

次に、 ξ が空間ベクトルの系列である場合を考える。 集合である場合と同様に $n' + 1$ 通りの特徴抽出を導出するが、各 f で抽出される特徴の次元数が異なる。

$$(2.8) \quad \begin{aligned} f_0(\xi) &= \{\langle \mathbf{p}_I \mathbf{p}_{I+1} \rangle\} \in \mathbb{R}^{m-1} \\ f_k(\xi) &= \{\langle \mathbf{p}_I \cdots \mathbf{p}_{I+k-1} \mathbf{e}_I^{-1} \rangle, I \in \mathcal{I}_k\} \in \mathbb{R}^{(m-k+1)|\mathcal{I}_k|} \end{aligned}$$

以上で、順序なし集合と順序つき集合に関する特徴抽出を系統的に導入した。 f_0 は内積であり、ある注目ベクトル \mathbf{p}_{l_1} の次の注目ベクトル \mathbf{p}_{l_2} に平行な成分の長さに $|\mathbf{p}_{l_2}|$ を掛けたものを示す。 f_1 は注目点の座標値を示す。 f_2 は2本のベクトルが各基底バイベクトル方向に作る面積を示す。 f_3 は3本のベクトルが各基底トライベクトル方向に作る体積を示す。 f_0 および f_n の特徴は計測座標系の回転に依存しない。

また、前述の Connell and Jain (2002) が提案した回転に不変な2種類の特徴は $|\mathbf{p}_{i+1} - \mathbf{p}_i| |\mathbf{p}_i - \mathbf{p}_{i-1}| = \gamma$ として、 $\cos(\theta) = \gamma(\mathbf{p}_{i+1} - \mathbf{p}_i) \cdot (\mathbf{p}_i - \mathbf{p}_{i-1}) = \gamma(\mathbf{p}_{i+1} \cdot \mathbf{p}_i - \mathbf{p}_{i+1} \cdot \mathbf{p}_{i-1} - \mathbf{p}_i \cdot \mathbf{p}_i + \mathbf{p}_i \cdot \mathbf{p}_{i-1}) \in f_0$ となり、 $\sin(\theta) = \gamma(\mathbf{p}_{i+1} - \mathbf{p}_i) \wedge (\mathbf{p}_i - \mathbf{p}_{i-1}) \mathbf{e}_{12}^{-1} = \gamma(\mathbf{p}_{i+1} \wedge \mathbf{p}_i - \mathbf{p}_{i+1} \wedge \mathbf{p}_{i-1} + \mathbf{p}_i \wedge \mathbf{p}_{i-1}) \mathbf{e}_{12}^{-1} \in f_2$ となる。 本研究で提案した空間ベクトルの集合の特徴に含まれることが分かる。

UCI レポジトリの手書き文字と手書き数字データから筆跡曲線上で等間隔に8つのポイント $\xi = \{\mathbf{p}_1, \dots, \mathbf{p}_8\}$ を取り出す。 そのポイントの系列から、 \mathcal{G}_2 を使用した提案手法により系統的に幾何的な特徴を抽出できる。 ここで、 $\mathbf{p}_i = x_i \mathbf{e}_1 + y_i \mathbf{e}_2$ であり、 $\sum_{l=1}^8 x_l = \sum_{l=1}^8 y_l = 0$ と数字の中心に原点を合わせてあるものとする。 図1に手書き数字‘7’の例を示し、図2に手書き文字‘H’と‘h’の例を示す。

特徴抽出 f_0 は、連続する2点の内積を抽出する。

$$(2.9) \quad \begin{aligned} f_0(\xi) &= [\langle \mathbf{p}_1 \mathbf{p}_2 \rangle, \dots, \langle \mathbf{p}_7 \mathbf{p}_8 \rangle] \\ &= [x_1 x_2 + y_1 y_2, \dots, x_7 x_8 + y_7 y_8] \in \mathbb{R}^7 \end{aligned}$$

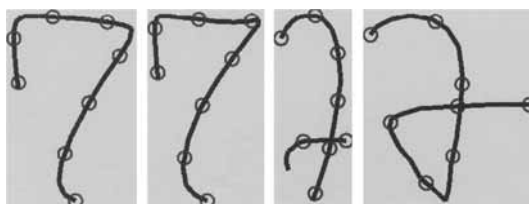


図1. Examples of hand-written digit ‘7’.

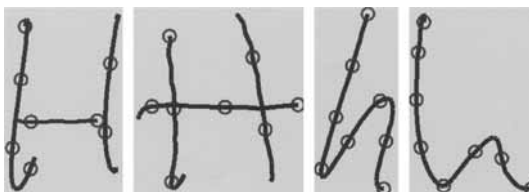


図2. Examples of hand-written character ‘H’.

特徴抽出 f_1 は、最も簡単で従来手法でもよく使われる特徴として、座標値を抽出する。

$$(2.10) \quad \begin{aligned} f_1(\xi) &= [\langle \mathbf{p}_1 \mathbf{e}_1^{-1} \rangle, \langle \mathbf{p}_1 \mathbf{e}_2^{-1} \rangle, \dots, \langle \mathbf{p}_8 \mathbf{e}_1^{-1} \rangle, \langle \mathbf{p}_8 \mathbf{e}_2^{-1} \rangle] \\ &= [\mathbf{p}_1 \cdot \mathbf{e}_1, \mathbf{p}_1 \cdot \mathbf{e}_2, \dots, \mathbf{p}_8 \cdot \mathbf{e}_1, \mathbf{p}_8 \cdot \mathbf{e}_2] \\ &= [x_1, y_1, \dots, x_8, y_8] \in \mathbb{R}^{16} \end{aligned}$$

特徴抽出 f_2 は、連続する2点の外積を抽出する。

$$(2.11) \quad \begin{aligned} f_2(\xi) &= [\langle \mathbf{p}_1 \mathbf{p}_2 \mathbf{e}_{12}^{-1} \rangle, \dots, \langle \mathbf{p}_7 \mathbf{p}_8 \mathbf{e}_{12}^{-1} \rangle] \\ &= [\mathbf{p}_1 \wedge \mathbf{p}_2, \dots, \mathbf{p}_7 \wedge \mathbf{p}_8] \\ &= [x_1 y_2 - x_2 y_1, \dots, x_7 y_8 - x_8 y_7] \in \mathbb{R}^7 \end{aligned}$$

2.3 多クラス分類問題に対する分布の学習

本研究では、前節で述べた提案手法によって抽出した特徴データについて、データ分布を学習し、多クラス分類を行う。Gaussian mixture model (GMM) は、データ空間におけるデータ分布を M 個の正規分布で近似する。ここで、 $M=1$ の場合は、2次判別に相当する。GMM はパラメータ $\Theta = \{\beta_j, \mu_j, \Sigma_j\}$ によって特徴づけられる。ここで、 β_j , μ_j および Σ_j はそれぞれ j 番目のガウス分布の混合比、平均および共分散行列である。GMM の出力は、

$$(2.12) \quad p(\xi | \Theta) = \sum_{j=1}^M \beta_j \mathcal{N}_d(f(\xi) - \mu_j; \Sigma_j)$$

で表される。ここで、 $\mathcal{N}_d(\cdot; \cdot)$ は平均が原点である d 次元ガウス分布関数を表す。

不完全データ $X = \{x_i = f(\xi_i) | 1 \leq i \leq N\}$ から、EM アルゴリズム (Dempster et al., 1977) により GMM のパラメータ Θ と潜在変数 $Z = \{z_{ij} \in \{0, 1\} | 1 \leq j \leq M\}$ を求める。ここで、 $z_{ij} = 1$ は各データ x_i が GMM の j 番目のガウス分布に属することを意味し、 $\sum_{j=1}^M P(z_{ij}) = 1$ を常に満たす。初期値として、 z_{ij} を適当に選択する。そして、 $P(Z)$ と Θ が収束するまで、E-ステップ (Expectation) と M-ステップ (Maximization) を繰り返し、潜在変数 Z とパラメータ Θ をそれぞれ更新する。E-ステップは、ベイズの定理 $P(Z | X, \Theta) \propto p(X | Z, \Theta)$ により、潜在変数 Z の確率を更新する。M-ステップは、尤度 $l(\Theta, X, Z) = p(X | Z, \Theta)$ を最大化するように、ガウス分布の変数 Θ を更新する。

しかし、GMM の大きな欠点は、自由パラメータ数が次元数の2乗 (d^2) と混合数 M に比例して増えるため、過学習が起こることである。自由パラメータを増やせば増やすほど、学習データの分布を正確に近似できるが、学習しないデータに対する分布から遠ざかってしまう。この欠点を緩和するために、Meinicke and Ritter (2001) が提案したように、各ガウス分布の共分散行列を固有値分解し、大きな固有値に対応する分散成分のみを用いることにする。不完全データ X に対する共分散行列の最大固有値 λ_{\max} と超パラメータ $\alpha \in (0, 1)$ を使用して、 $\lambda_- = \alpha \lambda_{\max}$ より大きな固有値に対応する成分のみを用いることにする。つまり、式(2.12)を展開し

$$(2.13) \quad \begin{aligned} p(x | \Theta) &= \sum_{j=1}^M \beta_j \prod_{k=1}^d \mathcal{N}_1((x - \mu_j) \cdot v_k; \lambda_k) \\ &\approx \sum_{j=1}^M \beta_j \left\{ \prod_{k=1}^{q_j} \mathcal{N}_1((x - \mu_j) \cdot v_k; \lambda_k) \right\} \\ &\quad \times \left\{ \mathcal{N}_1((x - \mu_j)_-; \lambda_-) \right\}^{d - q_j} \end{aligned}$$

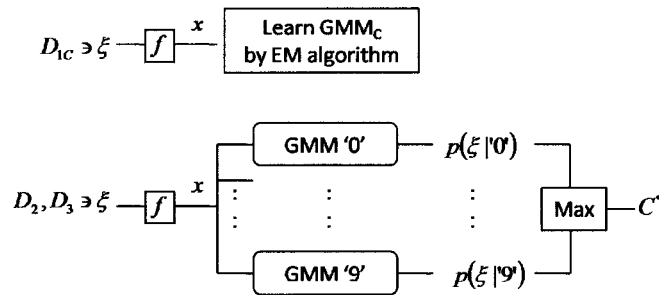


図 3. Flow of multi-class classification. The top diagram shows the training of the GMM for class $C \in \{ '0', \dots, '9' \}$. The D_{1C} denotes a subset of training samples whose label is C . The $f: \xi \mapsto x$ shows feature extraction. Either of $\{f_1, f_2, f_0\}$ is chosen as f . The bottom diagram shows estimation by the learned GMMs. The same f chosen for training is used here. The GMM_C outputs $p(\xi | C)$. The final estimation is $C^* = \arg \max_C p(\xi | C) P(C)$, where $P(C)$ is the prior distribution. The set D_3 consists of independent test data.

とする。ここで、 λ_k は j 番目のガウス分布の k 番目に大きな固有値、 v_k はそれに対応する固有ベクトルである。また、 $k \leq q_j \Leftrightarrow \lambda_k > \lambda_-$ であり、 $(x - \mu_j)_-$ は q_j 番目までの固有ベクトルと直交する成分の長さである。

多クラス分類問題の例として、手書き数字分類の流れを図 3 に示す。図 3 の上部の図に、数字 $C \in \{ '0', \dots, '9' \}$ 用 GMM の学習の流れを示す。 D_{1C} は数字 C の学習データであり、特徴抽出方法 $f: \xi \mapsto x$ は $\{f_0, f_1, f_2\}$ から選ばれる。図 3 の下部の図に、学習済み GMM による分類の流れを示す。分類時の f には学習時と同じ f を用いる。GMM $_C$ は確率密度 $p(\xi | C)$ を出力する。最後に $C^* = \arg \max_C p(\xi | C) P(C)$ により分類結果を出力する。ここで、 $P(C)$ は C の事前確率であり、 $\forall C, P(C) = \frac{1}{10}$ とした。 D_2 はバリデーション用データセットである。各 GMM $_C$ の混合数 M は D_2 に対する尤度により決定し、その後、決定した M を用いて、 D_2 に対する分類成功率により、 α を決定する。 D_3 はテストデータセットである。

2.4 エキスパートの混合

入出力関数の学習ではなく、データ分布の学習を行うことの大きなメリットに、弱学習機の推論を混合することによってより信頼できる推論を得られる点が挙げられる。GA を用いて導出する各特徴抽出は、それぞれ長所と短所を持つ。それらの長所を活かし短所を補うために、本研究では、次式の、推論の混合を用いて、分類を行う。

$$(2.14) \quad p(\xi | C) = \prod_{k=0}^2 p(f_k(\xi) | C)$$

各 GMM のパラメータには前節で述べた学習で獲得されたものを用いる。図 4 に、GA による特徴抽出 f_0, f_1, f_2 を通した推論の混合により数字を分類する流れを示す。

式(2.14)の事後確率の積のほかに、多数決も考えられるが、クラス数に比べエキスパートの数が少ないため、エキスパートによる識別結果が全て異なる場合が生じ、その場合は多数決できないという欠点がある。また、各特徴抽出が幾何的な情報を含んでおり、極端に識別率が低いエキスパートはないと期待できるため、本研究では式 2.14 によりエキスパートの識別結果を混合する。

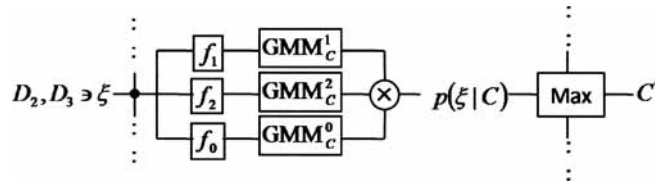


図 4. Mixture of GMMs. Three GMMs via different feature extractions are mixed to yield output $p(\xi|C)$.

3. 数値実験と考察

3.1 手書き文字(数字)分類問題

UCI レポジトリの Pen-Based Recognition of Handwritten Digits データセット (以下, Pendigits データセット) および UJI Pen Characters データセット (以下, Penchars データセット) を用いて多クラス分類の数値実験を行った。

Pendigits データセットは, 44 人が書いた 10992 サンプルから構成される。44 人のうち 30 人が書いた 7494 サンプルが学習データ D_1 とバリデーションデータ D_2 に分けられ, 残りの 14 人が書いた 3498 サンプルがテストデータ D_3 として用いられる。各サンプルを収集する際, 500×500 ピクセルの解像度を持つタブレット上で 100[msec] 間隔のペン先座標が獲得される。ペン先の軌道長を 7 等分する 8 ポイント $\{r_i\}$ が選ばれ, 縦方向の最大値が 100, 最小値が 0 となるようにスケーリングされる。本研究では, GA を用いる特徴抽出に先立ち, $p_i = r_i - \bar{r}$ と, 数字の中心を原点とした。

Penchars データセットは, 11 人が書いた 26 種類の文字(大文字と小文字を別々のクラスとしない)と 9 種類の数字の 1364 サンプルから構成される。クラスの合計が 35 クラスである。評価するには, 11 leaving-one-writer-out tests という方法を用いるテストを行った。つまり, 11 人が書いたデータのうち, 10 人が書いたデータを学習データとし, 残りの一人が書いたデータをテストデータとする 11 組の「学習・テスト」データセットが構成される。バリデーションデータは使用しないこととする。Pendigits データセットと同じように, ペン先の軌道長を 7 等分する 8 ポイント $\{r_i\}$ を選び, 縦方向の最大値が 100, 最小値が 0 となるようにスケーリングする。GA を用いる特徴抽出に先立ち, $p_i = r_i - \bar{r}$ と, 数字の中心を原点とした。

3.2 各特徴抽出と混合エキスパートでの分類結果

Pendigits データセットでは, 特徴抽出 $f \in \{f_0, f_1, f_2\}$ の各場合について, $\alpha \in \{0.1, 0.01, 0.001, 0.0001\}$ の 4 通りで D_1 を学習しそれぞれ D_2 に対する分類成功率を評価した。その結果を表 1 に示す。表中, 分類成功率の最大値に下線を引く。特徴抽出 f_0, f_1 について $\alpha = 0.01$ が選択され, 特徴抽出 f_2 について $\alpha = 0.0001$ が選択された。

表 1. Correct classification rate for D_2 of Pendigits.

α	f_0	f_1	f_2
0.1	89.89%	92.21%	95.52%
0.01	<u>94.40%</u>	<u>99.55%</u>	96.88%
0.001	93.76%	99.39%	96.83%
0.0001	94.00%	99.33%	<u>96.93%</u>

表 2. Identified models and number of misclassification via coordinates f_1 of Pendigits.

	'0'	'1'	'2'	'3'	'4'	'5'	'6'	'7'	'8'	'9'
M	4	9	6	3	4	7	2	8	9	7
\bar{q}	5.8	4.31	6.7	9	8.3	3.1	7.5	6.5	5.8	7.3
'0'	-								8	
'1'		-	3							
'2'		2	-							
'3'		2		-		2				
'4'					-	1				
'5'				4	1	-				9
'6'		1				1	-			
'7'		23		2				-		8
'8'									-	
'9'		1		1		19		2		-

表 3. Identified models and number of misclassification via outer products f_2 of Pendigits.

	'0'	'1'	'2'	'3'	'4'	'5'	'6'	'7'	'8'	'9'
M	1	5	3	3	4	5	1	5	7	6
\bar{q}	7	6.6	7	7	7	7	7	7	7	7
'0'	-						1			
'1'		-	17	1				1	1	1
'2'		1	-					4		
'3'	1	11		-		3			1	2
'4'					-					9
'5'				2	1	-			3	20
'6'	1				8		-		2	8
'7'		9	20	11				-	10	2
'8'						1			-	
'9'		1				14		8	1	-

表 2 に、特徴抽出 f_1 を用いた場合の各数字に対する GMM の混合数 M と平均有効次元数 $\bar{q} = \frac{\sum q_i}{M}$ および D_3 に対する誤分類数を示す。表 2 の誤分類数の見かたの例として、'0' の行と '8' の列のセルは、'0' であるデータのうち 8 サンプルが '8' として誤分類されたことを示す。誤分類の総数は 90 (分類成功率 97.42%) であった。同様に、表 3 に、外積特徴を抽出する f_2 を使用した場合の結果を示す。誤分類の総数は 176 (分類成功率 94.97%) であった。表 4 に、内積特徴を抽出する f_0 を使用した場合のモデル同定結果および分類結果を示す。誤分類の総数は 316 (分類成功率 90.97%) であった。

表 2, 3, 4 に示すように、特徴抽出によって、テストデータに対する分類結果が異なる。特徴抽出 f_1 では、各ポイントの座標情報が保たれるため、全般に誤分類が少ない。しかし脆弱な面もある。例えば '0' の学習データはほとんど全て数字の中心から見てタブレットの左上から書き始められていたため、テストデータで右上から書き始められる '0' に対し、分類に失敗した。 f_2 と f_0 では、各ポイントの座標情報が失われるが、数字の部分がどこに書かれたかに

表 4. Identified models and number of misclassification via inner products f_0 of Pendigits.

	'0'	'1'	'2'	'3'	'4'	'5'	'6'	'7'	'8'	'9'
M	4	9	3	3	3	6	2	9	8	9
\bar{q}	6	3.2	6.3	7	7	4.2	6	4.6	4	5.9
'0'	-								2	
'1'	1	-	21	11		13				2
'2'		12	-	24		5		6		
'3'		4	4	-		15				1
'4'		5			-	17		2		2
'5'	2	6	3	16		-			1	10
'6'	2		2			1	-	6	2	10
'7'	2	13	1	3	4		21	-	4	12
'8'		5				1		1	-	
'9'	2	8		4		25		2		-

表 5. Number of misclassifications by mixture of experts of Pendigits.

	'0'	'1'	'2'	'3'	'4'	'5'	'6'	'7'	'8'	'9'
'0'	-									
'1'		-	7							1
'2'			-	1						
'3'		2		-		1				
'4'					-					1
'5'				4		-				9
'6'	2					1	-			
'7'		8	1	2				-		15
'8'									-	
'9'		2		1		4		3		-

表 6. Correct classification rate of Penchars.

	f_0	f_1	f_2	Mixture of experts
Correct classification rate	46.70%	74.12%	51.39%	75.51%

関わらず部分的な形状を特徴として抽出する。そのため、 f_1 では分類に失敗した右上から書き始められた '0' も左上から書き始められた '0' と部分的な形状は同じであるので、'0' データに対する誤分類はほとんどなかった。

そこで、特徴抽出 f_1, f_2, f_0 のお互いの長所を活かすべく、混合エキスパートを利用した数字分類を行った。表 5 に混合エキスパートの分類結果を示す。誤分類の総数は 65 (分類成功率 98.14%) であり、 f_1, f_2 あるいは f_0 単独の分類結果より良くなった。

Penchars データセットでは、データ数が少ないためバリデーションを行わず、 $M=1$ とし、 $\alpha=0.01$ とする。

表 6 に示すように特徴抽出によって、混合エキスパートの分類結果は f_1 単独の分類結果より良くなり、分類成功率 75.51% を示す。

3.3 計測環境の揺らぎに対する頑健性

測定環境が揺らいだ場合を想定し、テストデータ D_3 の各数字をランダムに回転させたデータセット $D'_3 = \{R(\xi, \varphi), \xi \in D_3, \varphi \sim U(\varepsilon)\}$ を生成した。 $U(\varepsilon)$ は $[-\varepsilon, \varepsilon]$ の一様分布を示し φ は確率変数である。 $R(\xi, \theta)$ は ξ に属する全ポイントを一律に角度 θ だけ回転させる。 $\varepsilon \in \{\frac{\pi}{20}, \frac{\pi}{10}, \frac{\pi}{5}\}$ でそれぞれ 20 回 D'_3 を生成し分類テストを行った。

特徴抽出 f_1 を用いた場合と混合エキスパートを用いた場合の、Pendigits データセットと Penchars データセットの分類成功率の平均と標準偏差を図 5 と図 6 にそれぞれ示す。特徴抽出 f_1 のみを用いた場合の分類精度は ε が大きくなるにつれ著しく低下したが、混合エキスパートを利用した分類精度はそれほど低下しなかった。なお、 f_0, f_2 については、それらで抽出する

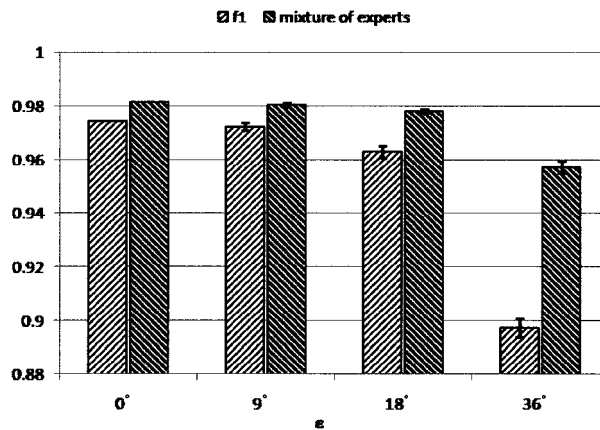


図 5. Correct classification rate with f_1 and mixture of experts of Pendigits.

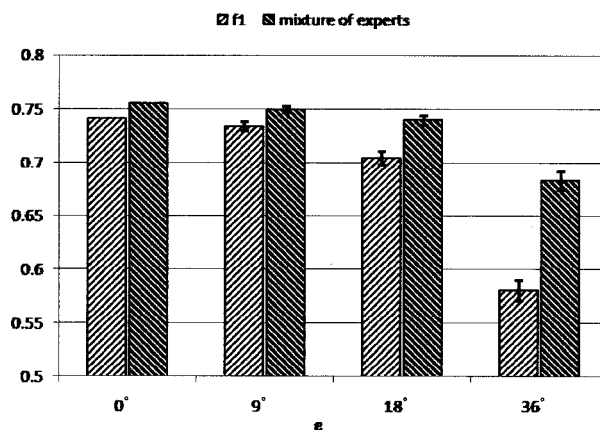


図 6. Correct classification rate with f_1 and mixture of experts of Penchars.

特徴は座標を回転しても変わらないため、どんな ε で作成した D'_3 に対しても分類結果は同じである。

4. まとめ

本研究では、GA を用いる系統的な特徴抽出手法を提案した。抽出した特徴を使って、混合正規分布を用いた手書き数字分類問題に適用した。手書き数字のような 2次元の対象に提案手法を適用した場合、座標値と外積および内積の 3通りの特徴抽出が導出される。座標値特徴と外積特徴および内積特徴それぞれ単独で分類したとき、テストデータに対して、座標値特徴を用いた場合の分類成功率が一番高かった。しかし、異なる測定環境を想定して座標を回転変換したテストデータに対しては、座標値特徴による分類成功率は低下した。それに対し、GA で系統的に導出した特徴抽出の混合エキスパートを用いると、測定環境が学習時とテスト時で同じであっても分類成功率は高く、また、テスト時の測定環境の揺らぎを大きくしても分類精度はそれほど低下しなかった。したがって、従来よく用いられる座標値特徴と比べ、GA を用い導出する複数の特徴抽出を混合する戦略は、分類精度および頑健性において優越であることが確認できた。

参 考 文 献

- Abo-Zaid, A., Hinton, O. and Horne, E. (1988). About moment normalization and complex moment descriptors, *Proceeding of the 4th International Conference on Pattern Recognition*, 399–407.
- Asuncion, A. and Newman, D. J. (2007). UCI Machine Learning Repository, University of California, School of Information and Computer Science, Irvine, California.
- Bahlmann, C. and Burkhardt, H. (2004). The writer independent on-line handwriting recognition system frog on hand and cluster generative statistical dynamic time warping, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **26** (3), 299–310.
- Buchholz, S. and Bihan, N. L. (2006). Optimal separation of polarized signals by quaternionic neural networks, *14th European Signal Processing Conference, EUSIPCO*, Florence, Italy.
- Connell, S. D. and Jain, A. K. (2002). Writer adaptation for online handwriting recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **24** (3), 329–346.
- Dempster, A., Laird, N. and Rubin, D. (1977). Maximum likelihood from incomplete data via the EM algorithm, *Journal of the Royal Statistical Society Series B*, **39** (1), 1–38.
- Doran, C. and Lasenby, A. (2003). *Geometric Algebra for Physicists*, Cambridge University Press, Cambridge.
- Dorst, L., Fontijne, D. and Mann, S. (2007). *Geometric Algebra for Computer Science: An Object-oriented Approach to Geometry*, Series in Computer Graphics, Morgan Kaufmann, San Francisco, California.
- Guyon, I., Albrecht, P., LeCun, Y., Denker, J. and Hubbard, W. (1991). Design of a neural network character recognizer for a touch terminal, *Pattern Recognition*, **24**, 105–119.
- Hestenes, D. (1999). *New Foundations for Classical Mechanics*, 2nd ed., Kluwer Academic, Dordrecht.
- Hildenbrand, D. and Hitser, E. (2008). Analysis of point clouds using conformal geometric algebra, *3rd International Conference on Computer Graphics Theory and Applications*, Funchal, Machira, Portugal.
- Hirose, A. (2003). *Complex-valued Neural Networks: Theories and Applications*, Series on Innovative Intelligence, World Scientific, Singapore.

- Hitzer, E. (2007). Quaternion Fourier transform on quaternion fields and generalizations, *Advances in Applied Clifford Algebras*, **17**(3), 497–517.
- Hu, J. Y., Lim, S. G. and Brown, M. K. (1999). Writer independent on-line handwriting recognition using an HMM approach, *Pattern Recognition*, **33**, 133–147.
- Hu, M. K. (1962). Visual pattern recognition by moment invariants, *IRE Transactions on Information Theory*, **8**(2), 179–187.
- Matsui, N., Isokawa, T., Kusamichi, H., Peper, F. and Nishimura, H. (2004). Quaternion neural network with geometrical operators, *Journal of Intelligent and Fuzzy Systems*, **15**(3–4), 149–164.
- Meinicke, P. and Ritter, H. (2001). Resolution-based complexity control for Gaussian mixture models, *Neural Computation*, **13**(2), 453–475.
- Nitta, T. (1997). An extension of the back-propagation algorithm to complex numbers, *Neural Networks*, **10**(8), 1391–1415.
- Sekita, I., Kurita, T. and Otsu, N. (1992). Complex autoregressive model for shape recognition, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, **14**(4), 489–496.
- Sommer, G. (2001). *Geometric Computing with Clifford Algebras*, Springer-Verlag, Heidelberg.

Feature Extraction by Geometric Algebra from Geometric Data

Minh Tuan Pham¹, Kanta Tachibana¹, Eckhard Hitzer², Sven Buchholz³,
Tomohiro Yoshikawa¹ and Takeshi Furuhashi¹

¹School of Engineering, Nagoya University

²School of Engineering, University of Fukui

³Institute for Informatics, University of Kiel

Most conventional methods of feature extraction for pattern recognition do not pay sufficient attention to inherent geometric properties of data, even where the data have characteristic spatial features. In this study, we introduce geometric algebra to systematically extract invariant geometric features from spatial data given in a vector space. Geometric algebra is a multidimensional generalization of complex numbers and of quaternions, and can accurately describe oriented spatial objects and relations between them. We further propose a combination of several geometric features using Gaussian mixture models. We demonstrate our new method by classification of hand-written digits and alphabetic characters.

Key words: Geometric algebra, feature extraction, Gaussian mixture model, pattern recognition, mixture of experts.