

氏 名 田中 英希

学位（専攻分野） 博士（統計科学）

学位記番号 総研大甲第 1236 号

学位授与の日付 平成 21 年 3 月 24 日

学位授与の要件 複合科学研究科 統計科学専攻  
学位規則第 6 条第 1 項該当

学位論文題目 Statistical Analysis of 2D Partial Shapes Using  
Protrusion Fourier Descriptor

論文審査委員 主 査 教授 種村 正美  
教授 田村 義保  
教授 尾形 良彦  
教授 城石 俊彦

## 論文内容の要旨

世界を視覚的に知覚し理解するための重要な概念である「形」は、本質的に多次元の特質である。それゆえ形は定量的に評価するのが難しく、自然科学において長い間研究対象ではなかった。本論文において扱われる形とは、2次元空間における物体の内側と外側の境界線の形状であり、したがって、それ自身と交差しない一筆書きの閉曲線または開曲線の形状である。多くの場面で、一筆書きの曲線形状を記述するのに用いられてきた手法の一つにフーリエ記述子(FD)がある。この手法では、まず曲線上の各点に対し、その点における何らかの幾何学的情報が数値として対応づけられる。そして始点から終点に向けて並べられたその数値の列の離散フーリエ変換として図形は記述される。各点に対応づけられる幾何学的情報が異なるさまざまなFDが開発され、ある種のFDについて、不規則形状に対する優れた特徴抽出能力をもつこと、不規則形状の統計解析において有用であること、などが示されてきた。本論文では曲線形状の記述方法としてFDに焦点をあてる。

部分形状の統計解析で、部分形状の数値記述に用いられてきた開曲線記述子に接線フーリエ記述子(TFD、P型フーリエ記述子)がある。TFDは、浮世絵の作者判別、ヒトの横顔輪郭の認識、ハナハスの花卉やイネの葉の形状の統計解析などに用いられ、それぞれの場面での有用性が示されてきたが、TFDは、曲線上の隣接点間の位置の差分系列のフーリエ変換であり、ノイズを増幅する傾向がある。そこで、不規則で複雑な開曲線形状の微妙な特徴をとらえる能力について、さらに優れた記述方法を求めて、protrusion フーリエ記述子(PFD)を考案した。PFD法では、始点から終点への向きに対し左右からのprotrusion(突出度合い)が、曲線上の各点において求められる。左右からのprotrusionをそれぞれ実数、虚数成分とする複素数が、曲線上の各点に対応付けられ、PFDは、始点から終点にむけて順序づけられたその複素数列のフーリエ変換として定義される。ちなみに曲線上の各点におけるprotrusionは、曲線形状に対し一意に定まるスケルトン・垂線グラフと名付けられた場を用いて計算される。

さらに不規則形状に対する特徴抽出能力に関しTFDとPFDを比較する実験を行った。用いたデータは、9つの系統(合計93匹)のマウスの下顎骨から特定部位を表す93個の開曲線であり、まず、TFD、PFDそれぞれを用いて曲線を数値表現した。系統の判別モデル(ロジスティック回帰モデル)を当てはめ、モデルの良さを一般化情報量規準(GIC)を用いて情報論的に測った。もし用いた記述子が系統ごとの形状の差をより良くとらえるならば、データは系統ごとに明確に分離し、データに当てはめられた判別モデルのGIC値はより低くなるはずである。結果として、TFD使用時よりPFD使用時の方がGIC値は低かった。これは、PFD使用時の方がより正確に系統を判別できること、つまり、PFDがTFDより不規則形状に対する特徴抽出能力において優れていることを示している。

次に、2つの系統を起源とする F2 世代（合計 207 匹）のマウスを用い下顎骨の部分形状の遺伝解析を行った。本解析の目的は形態形成に対して影響力の大きい遺伝子の探索と、各遺伝子の影響の様子の視覚化による解釈である。全 207 個体について、ゲノム上の特定の 122 箇所において、分子マーカーを用いて遺伝型が調べられている。形は量的形質であり、一般に量的形質は、相互作用する複数の遺伝子の影響を受けている。最も単純な場合として 2 つの遺伝子の相互作用（エピスタシス）を考慮し、ゲノム上の位置（の付近にある遺伝子）の任意のペアについて、その下顎骨形状に対する影響力を測った。また、その影響力の測定方法として、本質的に多変量の特質である形を、多変量のまま解析するために、GIC 値を検定統計量として用いる仮説検定を採用した。まず形のデータを、遺伝型の情報を用いずにランダムに分割し、それに判別モデルをあてはめ、モデルの良さを GIC 値で測る、という計算を十分な回数くりかえし、GIC 値の null 分布（形にその遺伝型が何の影響力ももたない、という仮説の下での null 分布）を用意しておいた。ゲノム上の任意の位置のペアについて、そのペアの遺伝型の情報を用いて形のデータを分割し、判別モデルを当てはめ、モデルの良さを GIC 値で測る。そして、その GIC 値の null 分布に対する片側 P 値を求め、それをその遺伝子の形に対する影響力とみなした。本解析の結果、異なる部位に対し、形に対して大きな影響力をもつゲノム上の位置のペアが大きく異なること、さらに記述子からの曲線形状の視覚化により、各々の位置のペアの形に対する影響の仕方が、それぞれ異なることなどが分かった。

## 論文の審査結果の要旨

申請者の博士申請論文の主要成果は、2次元形状に関して開曲線の新しい数値表現法を提案して、従来の方法に比べて優れた手法であることを実証的に示していることである。

論文は全70頁で6章と1つのAppendixからなり、英語で執筆されている。本研究の主たる目的は、形状の統計解析を行うことである。特に開曲線の数値化に注目し、Fourier記述子を新たに提案している。提案した記述子を用いて、マウスの下顎骨の形状と遺伝子情報との関連を探っている。論文の構成は次のとおりである。第1章は序章である。第2章では閉曲線のための既存のフーリエ記述子、complex フーリエ記述子(CFD)、楕円フーリエ記述子(EFD)について詳解している。第3章では開曲線のための既存のフーリエ記述子であるtangent フーリエ記述子(TFD、またはP型フーリエ記述子)について説明した後、新しい開曲線記述子、Protrusion フーリエ記述子(PFD)を提案している。詳細な計算アルゴリズム、および逆変換アルゴリズムは付録に記載し、幾何学的意味と計算手順の概略を説明している。第4章では、TFDとPFDの記述力の差異を、「9つの近交系の合計93匹のマウスの下顎骨の特定部位の形状を表す93個の開曲線」を例として検討している。主成分分析を用いて得た $K(\leq 93)$ 個の主成分得点に対して多項ロジスティック回帰モデルを考え、あてはまりの良さを、一般化情報量規準(GIC)を用いて比較している。その結果、PFDを用いて曲線を記述した場合の方が、TFDを用いた場合よりも、GICは小さい値をとることを示している。このことより、PFD法が既存のTFD法よりも、不規則な部分形状の統計解析に用いる開曲線記述子として適していると結論している。第5章においては、PFDを用いた形の遺伝子解析を行っている。2つの近交系の交配実験により作出したF2世代の合計207個体のマウスの下顎骨のある特定部位の形状をPFD法により数値化した後、主成分分析を行うことにより求めた主成分得点を用いて、二項ロジスティック回帰モデルに対するGIC値を求めることにより、形に対して影響力をもつ遺伝子の交互作用(エピスタシス)について検出し、それを視覚化により解釈できることを示している。第6章では結論と将来の発展が述べられている。付録は輪郭線からの開曲線の抽出法、PFDの計算アルゴリズム及びPFDから曲線形状への逆変換法の説明にあてられている。

形状を数値化し、その数値と対象物の特性との関係を求めることは、多くの領域で必要とされている。本論文では、2次元形状に限られてはいるが、開曲線の新しい数値表現を提案し、形状が複雑な場合に従来の方法と比べて優位性を持つ手法であることは高く評価できる。また、提案する方法を用いることにより、形状を決める遺伝子のエピスタシス解析が可能であることを示したことも評価できる。

なお、本研究の主要部分であるPFDに関して、申請者を筆頭著者とする査読付き論文がすでに英文誌に掲載されている。

以上により、申請者の博士申請論文は博士(統計科学)の学位を授与するのに十分な内容をもつものと判断された。