

## 健康状態の推定を目的とした排便物識別のための一手法

小松原 美穂\* 石川 千里\* 高田 雅美\* 鴨 浩靖† 新出 尚之† 城 和貴\*

*miho@ics.nara-wu.ac.jp*

\* 奈良女子大学 大学院人間文化研究科 情報科学専攻

† 奈良女子大学 理学部 情報科学科

### 概要

本論文において、便の形状を分類するシステムを提案する。このシステムには、チェインコードと楕円フーリエ記述子を用いる。再現率と誤認識率を調べることによって、提案システムの評価を行う。実験の結果、再現率は75%以上となった。

## Auto Classification of Feces for Health Condition

Miho Komatsubara\* Chisato Ishikawa\* Masami Takata\* Hiroyasu Kamo† Naoyuki Nide† Kazuki Joe\*

\* Graduate School of Humanities and Sciences, Nara Women's University

† Faculty of Science, Nara Women's University

### Abstract

In this paper, we propose a system to classify shape of feces. In the proposed system, a chain code and elliptic Fourier descriptors are applied. To evaluate the system, a recall and a misidentification rates are calculated, and the recall rate is emphasized in this paper. The effectiveness of the proposed system is established. As the result, each recall rate becomes more than 75%.

## 1 はじめに

人の健康状態を知るバロメータとして、血液や血圧、心拍数等が考えられる。これらを測定するためには、通院や器具の装着が必要となる。そのため、人に対して心身ともに負担となる。そこで、利用者に負担をかけず、自動的に健康状態を判定するためのシステムの開発が望まれている。我々は、そのためのバロメータとして便に着目する。便は、健康や食生活と関係が深く、日々の健康状態を把握する一番簡単なバロメータである [1]。本研究では、自動撮影された便画像を認識・分類するシステムを開発する。なお、便の分類には、Bristol Stool Form Scale [2] を用いる。

以下、2章では、健康状態を知るためのバロメータについて述べる。3章において、便画像の自動分類システムを提案する。提案システムを評価するために、4章において実験を行う。

## 2 健康状態を知るための手段

近年、健康で長生きをするために、自分の状態を把握することが望まれている。そのため、人の健康状態を知るバロメータが必要である。バロメータとしては、血液・血圧・心拍・便などがある。

血液・血圧・心拍を取得するためには、意識的に心

かけての通院や器具の装着が必要となる。このことは、利用者にとって心身ともに負担となる。特に、血液から健康状態を知るためには、採血をしなければならないため、毎日行うことによって健康を害する可能性がある。

一方、便は、生理的現象であるために、身体への負担を回避することが可能である。また、排便の際に自動的に便に関する情報を取得することが可能となれば、利用者の負担をなくすることができる。

便に関する情報を取得するための一般的な方法として、検便が行われている。しかし、この方法は、自動的に取得することが難しく、利用者への負担が生じる。また、検便した便を検査するためには、専用の設備が必要となる。そこで、便の情報を自動的に取得するためには、検便ではなく、便画像から健康状態を把握できるようにシステムを開発すべきである。

便画像には、主に、形状・大きさ・色彩の3情報が含まれる。これらをそれぞれ独立に比較し、比較情報を統合することによって、健康状態を把握することが可能であると考えられる。そこで、本論文では、まず、形状情報を用いて分類するためのシステムを開発する。便の形状の分類尺度として、表1の Bristol Stool Form Scale を適用する。これらの Type のうち、Type6 と7に関しては、明確な形状が存在しないため、本論文の形状分類システムからは省く。

表 1: The Bristol Stool Form Scale

Type	Explanation
1	Separate hard lumps like nuts (difficult to pass)
2	Sausage shaped but lumpy
3	Like a sausage but with cracks on surface
4	Like a sausage or snake, smooth and soft
5	Soft blobs with clear-cut edges (passed easily)
6	Fluffy pieces with ragged edges, a mushy stool
7	Watery, no solid pieces (entirely liquid)



図 1: サンプル画像の一例

### 3 自動分類システムの提案

本章では、便を自動的に取得し形状分類するためのシステムを提案する。実際の便の形状は、複数の Type の特徴を持つ可能性がある。そこで、本システムでは、1つの Type に断定せず、複数の Type に分類されることを許す。提案システムは、次の 3 つの過程で行われる。

1 つ目の過程では、前処理として便画像のデータを加工する。取得された画像には、単体の便、複数で重なっていない便、複数で重なり合う便の 3 種類がある。そこで、まず、背景画像から便を切り出す。その際、取得された画像に複数の便が含まれている場合、1 つずつの便に切り分ける処理を行う。次に、各 1 つの便を白い画像に貼り付け、2 値化を行う。

2 つ目の過程として、画像の特徴量を計算する。特徴量を得るために、形状情報を定量化する必要がある。本提案では、相似変換、回転、平行移動、計測開始点に対して不変となる定量化手法の 1 つである楕円フーリエ記述子を適用する。そのため、特徴量は楕円フー

図 2: テスト画像の一例 (1 型)

表 2: テスト画像とサンプル画像の類似度

ID	Type				
	1	2	3	4	5
1	5.49E-03	2.06E-01	2.20E-01	4.53E-01	1.09E-01
2	1.91E-02	2.53E-01	3.74E-01	5.26E-01	1.05E-01
3	1.96E-02	2.78E-01	3.20E-01	4.53E-01	5.68E-02
4	9.41E-03	2.74E-01	3.48E-01	4.32E-01	1.07E-01
5	2.81E-02	2.60E-01	2.99E-01	4.38E-01	9.52E-02
6	1.02E-02	2.24E-01	2.99E-01	4.18E-01	1.25E-01
7	1.47E-02	2.10E-01	3.07E-01	4.36E-01	7.43E-02
8	2.22E-02	3.02E-01	3.39E-01	4.93E-01	2.01E-01
9	1.74E-02	2.27E-01	3.12E-01	4.72E-01	6.82E-02
10	3.15E-02	3.19E-01	3.14E-01	4.61E-01	1.19E-01
11	7.87E-03	2.64E-01	2.80E-01	4.36E-01	9.68E-02
12	4.54E-03	2.61E-01	3.29E-01	4.45E-01	1.01E-01
13	1.26E-02	2.70E-01	3.35E-01	4.73E-01	6.24E-02
14	2.85E-02	2.48E-01	2.98E-01	4.86E-01	6.38E-02
15	1.34E-02	2.92E-01	2.98E-01	4.37E-01	6.11E-02
16	1.95E-02	2.20E-01	2.58E-01	4.37E-01	5.93E-02
17	8.42E-03	2.51E-01	2.94E-01	4.92E-01	5.66E-02
18	2.41E-02	3.00E-01	3.43E-01	4.91E-01	1.37E-01
19	1.44E-02	1.78E-01	3.04E-01	4.58E-01	2.26E-01
20	1.20E-02	2.73E-01	2.91E-01	4.50E-01	5.58E-02

エ係数で表される。楕円フーリエ記述子は、チェーンコードで表した形状を離散フーリエ変換することで算出される [3]。チェーンコードは、効果的な曲線の記述法であり、数列によって輪郭線を記述することが可能である。楕円フーリエ記述子の係数は、対象物の向きや大きさ、計測開始点に関して不変となるように、第 1 調和の楕円を基に、数学的に標準化する必要がある。

3 つ目の過程として、特徴量を基に分類処理を行う。分類処理のためのアルゴリズムは、以下ようになる。

1. 前処理を施した便画像の読み込み
2. 便の特徴量を計算
3. サンプル画像の特徴量と Step2 の特徴量間の類似度を計算
4. 類似度と Bristol Stool Form Scale の各 Type の閾値を比較
5. 閾値より小さい場合、その Type に分類

ここで、類似度は特徴量の差の 2 乗和とする。図 1 は、サンプル画像の 1 例である。左から、Type 1, 2, 3, 4, 5 である。サンプル画像では、各 Type の便を 1 枚の画像に表記し、それを  $N$  枚用意する。また、各 Type の閾値は、サンプル画像とテスト画像の特徴量の最小類似度を計算し、最も小さい最小類似度を持つ Type と 2 番目に小さい最小類似度を持つ Type を比較することによって決定する。たとえば、図 2 のような Type1 のテスト画像に対し、 $N = 20$  枚のサンプル画像を比較した場合、類似度は、表 2 となる。ここで、太字の値は、各 Type における最小類似度である。表 2 より、Type1 の最小類似度が最も小さく、Type5 が 2 番目に小さい値である。図 3 は、Type1 の  $N$  個の

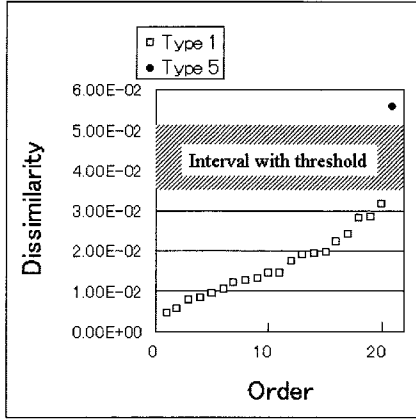


図 3: 閾値の範囲 (1 型)

類似度と Type5 の最小類似度をソートした結果である。ここで、縦軸は類似度、横軸は順位を表す。また、Type1 のデータは四角、Type5 のデータは円のマークで示す。図 3 より、20 位の Type1 の類似度の値と、21 位の Type5 の類似度の値の間の斜線部に閾値が存在すると目安をつけることができる。本システムでは、斜線部の幅に影響されることなく、閾値は 2 番目に小さい最小類似度と定義する。

## 4 実験

本論文で提案した便画像の形状分類システムの性能を調べるために、実験を行う。実験には、CPU が AMD Athlon(tm) XP 2500+ (1833MHz)、L1 キャッシュが 64KB、L2 キャッシュが 64KB、L3 キャッシュが 512KB、メモリが 256MB の計算機を用いる。チェインコードおよび楕円フーリエ係数までの計算には SHAPE を用いる [3]。SHAPE は、Windows 上で動作し、楕円フーリエ記述子の導出や主成分分析、輪郭再構築等の画像解析機能を持つフリーのソフトウェアである。

本実験では、目視分類は正しいと仮定し、その分類がどれだけ再現されているかによってシステムの性能を評価する。評価に用いる実験データとして、38 枚の便画像を用意する。比較する分類方法として、2 種類用いる。1 つは、「形状・大きさ・色彩情報を用いた目視分類」、もう 1 つは、「形状情報のみを用いた目視分類」を用い、以下、それぞれを「分類 1」「分類 2」と定義する。

システムの性能を示す  $P(i, j)$  は、次の式により求

表 3: 分類 1 による分類結果

Type	1	2	3	4	5
Number	19	3	6	8	2

表 4: 分類 1 に対する  $P(X_i, Y_j)$

$i \setminus j$	1	2	3	4	5
1	<b>1.00</b>	0.33	0.16	0.00	1.00
2	0.00	<b>0.33</b>	0.33	0.12	0.00
3	0.00	0.00	<b>0.16</b>	0.00	0.00
4	0.00	0.00	0.16	<b>0.37</b>	0.00
5	0.78	1.00	0.66	0.37	<b>1.00</b>

められる。

$$P(i, j) = \frac{N(Y_i \cap X_j)}{N(X_j)}, \quad (1)$$

ここで、 $X_i$  は  $i$  型に目視分類されたことを、 $Y_j$  は  $j$  型に自動分類されたことをそれぞれ示す。また、 $N(I)$  は  $I$  を満たす画像の数である。 $X_i = Y_j$  のとき、 $P(X_i, Y_j)$  は、 $i$  型に関する目視の再現率となる。そのため、システムの性能が高いとは、再現率が高いことを意味する。また、 $P(X_i, Y_j)$  は、 $i$  型と目視分類された事例に対する  $j$  型への誤認識率とする。提案システムでは、複数の型への分類を認めているため、誤認識率ではなく、再現率が重要である。ただし、再現率が低い場合は、正しい分類が行われていないと判断する。

以下、4.1、4.2 節では、再現率を用いてシステムの評価を行う。

### 4.1 分類 1 を用いた評価

本節では、 $X_i$  を分類 1 とした場合の  $P(X_i, Y_i)$  を計算し、システムの性能を調べる。38 枚の実験データを分類 1 によって分類した結果を表 3 に示す。表 3 において、Type5 の画像は 2 枚のみである。そのため、Type5 に関してシステム性能を正しく評価できない。そこで、Type5 の再現率は参考にしなすことにする。

分類 1 とシステムの自動分類との比較結果  $P$  を、表 4 に示す。各 Type に関する再現率は太字で示される。比較結果より、Type1 の再現率が 100% であるため、分類が成功している。本システムは形状情報のみを用いて自動分類を行うため、Type1 は形状情報のみで分類可能であると考えられる。一方、Type2, 3, 4 の再現率は低く、正しい分類が行われていない。これは、分類 1 において、形状に加え、大きさや色彩情報も混在しているためであると考えられる。よって、Type2, 3, 4 の便を分類するためには便画像が持つ全情報を組み合わせるべきである。

表 5: 分類 2 による分類結果

Type	1	2	3	4	5
Number	9	0	1	4	24

表 6: 分類 2 に関する  $P(X_i, Y_j)$

$i \setminus j$	1	4	5
1	<b>1.00</b>	0.00	0.58
2	0.00	0.25	0.12
3	0.00	0.25	0.00
4	0.00	<b>0.75</b>	0.40
5	0.55	0.00	<b>0.91</b>

表 4 より, Type1 の Type5 に対する誤認識率は 100% である. これは, Type1 と Type5 が似た形状を持つためであると考えられる.

以上より, 本システムを用いて分類 1 を再現することは難しい. ゆえに, 今後の課題として, 大きさや色彩情報に関する分類システムも必要である.

## 4.2 分類 2 を用いた評価

4.1 節で用いた分類 1 は形状・大きさ・色彩情報を用いて分類されている. 一方, 本システムでは, 形状情報を用いて分類を行う. そのため, 本システムを用いた形状分類の性能を正しく判断することが難しい. そこで, 本節では, 形状のみの性能を評価するために, 分類 2 を用いて実験を行う.

実験データに対して分類 2 を行った結果を表 5 に示す. 表 5 より, 分類 2 で Type2, 3 に分類された画像は, それぞれ 0, 1 枚のみである. そのため, システムの性能を正しく評価するために, Type2, 3 に関する実験結果は考慮しない.

Type1, 4, 5 に関する比較の結果を表 6 に示す. 表 6 より, Type1, 4, 5 のそれぞれの再現率が 100%, 75%, 91% と高いことが分かる. この結果から, 形状情報判断のみに関しては, 本システムを用いることで目視と同等の判断ができると考えられる. また, 表 6 より, Type1, 4, 5 それぞれの誤認識率は再現率と比べて低いことが分かる. ゆえに, 誤分類は, 我々のシステムを用いることによって防ぐことができる.

以上より, 本システムは形状情報を正しく認識し, 分類することが可能であると考えられる.

## 5 まとめ

本論文では, 便画像を自動認識し, 便の形状を評価・分類するシステムを提案した. 本システムには, 形状情報を評価する手法の 1 つである楕円フーリエ記述子を用いた. また, サンプル画像とテスト画像を用いて, Bristol Stool Form Scale における Type1 から Type5 それぞれに対し, 固定の閾値を設定し, その閾値を用いて便の Type の候補を選出させた.

提案システムの性能を調べるため, 目視分類した便と本システムによって自動分類した便の Type を比較し, 再現率を求める実験を行った. 4.1 節の実験の結果, Type1 および Type5 に関しては, 再現率が 100% となったことから, 分類が成功していることが分かった. しかし, Type2, 3, 4 においては, 再現率が低く, 正しい分類が行われなかった. この原因として, 実験データの目視分類が, 形状だけではなく, 大きさや色彩の情報も混在していることが考えられる. また, 本システムの性能を測るために, 4.2 節にて形状情報の判断のみによる目視分類が再現可能かどうかを確認した. その結果, Type1, 4, 5 において, 再現率が高いことが分かった. 以上の実験から, 形状情報判断のみに関しては, 本システムを用いることで目視と同等の判断ができること, 目視で便を分類する際には色情報・大きさ情報にも大きく依存することが分かった.

今後の課題として以下の点があげられる. 1 つ目は, 再現率を上げることである. そのために, サンプル画像とテスト画像に実際の便画像を用いることで, より精度の高い閾値を与えることが出来ると考えられる. 2 つ目は, より正確な結果を得るために, 実験データの枚数を増やすことである. 3 つ目として, 本システムに対して, 大きさと色彩情報に関する分類機能を加え, より正確な自動分類を行えるように改良することがあげられる.

## 謝辞

本研究の一部は NEC システムテクノロジーによる.

## 参考文献

- [1] George Longstreth, Grant Thompson, William Chey. et al. *Functional bowel disorders*. *Gastroenterology*, 130:1480-1491, 2006.
- [2] Luke O'Donnell, Jim Virjee, Kenneth Heaton. *Detection of pseudodiarrhoea by simple clinical assessment of intestinal transit rate*. *British Medical Journal*, 300(6722):439-440, Feb 1990.
- [3] Hiroyoshi Iwata, Yasuo Ukai. *SHPAE: A computer program package for quantitative evaluation of biological shapes based on elliptic Fourier descriptors*. *Journal of Heredity*, 93(5):384-385, Sep 2002.