

## 細胞領域分割のための画像処理 GP における学習領域決定法の検討

## Decision Method of Learning Region in Image-Processing GP for Cell Segmentation

関谷駿介<sup>\*1</sup> 布川将来人<sup>\*1</sup> 小泉範子<sup>\*2</sup> 奥村直毅<sup>\*2</sup> 山本詩子<sup>\*2</sup> 廣安知之<sup>\*2</sup>  
 Shunsuke Sekiya Sakito Nunokawa Noriko Koizumi Naoki Okumura Utako Yamamoto Tomoyuki Hiroyasu

<sup>\*1</sup>同志社大学大学院生命医科学研究科  
 Graduate School of Life and Medical Sciences Doshisha University

<sup>\*2</sup>同志社大学生命医科学部  
 Faculty of Life and Medical Sciences Doshisha University

In tissue engineering for the corneal endothelium, the quality evaluation of cultured cell is performed by extracting the feature values from cell images. In order to support the evaluation, we previously proposed a feature extraction system that uses learning region and Genetic Programming. In this paper, a decision method of the learning region is proposed by using texture feature values and clustering. The experiment shows that the proposed method improves the performance of the system.

## 1. はじめに

近年、コンタクトレンズや眼内手術の普及、社会全体の高齢化により角膜障害が増加している。この角膜障害を解決するための新たな治療法として、角膜再生医療が注目されている [Koizumi 09]。角膜再生医療では、患者やドナーの眼球から得た角膜内皮細胞を体外で培養し、培養した細胞を患者に移植することによって治療を行う。品質評価の方法として、光学顕微鏡を用いて撮像した角膜内皮細胞画像から細胞の密度や面積のばらつき、形状といった特徴量を計測する方法が挙げられる。現状の角膜再生医療では、専門家がこれらの特徴量を目視で判断することにより、細胞の品質評価を行っている。一方で、角膜内皮再生医療をさらに普及させるためには、細胞の培養を自動化することが必要であり、細胞の品質評価も自動化しなければならない。品質評価を自動化するためには、専門家が判断するのではなく細胞画像から個々の細胞領域を分割する画像処理を行い、特徴量を計測する必要がある。現在、特徴量を画像処理により解析する様々なソフトウェアが開発されている。しかしながら、これらのソフトウェアでは対象とする画像毎にユーザーが複雑な画像処理フィルタの組み合わせや閾値を決定しなければならない。そのため、ユーザーに画像処理の知識が必要であり、多量の細胞画像に対して品質評価を行う際には多くの時間と労力が必要となる。

これらの問題を解決するために、我々は過去に進化的計算の一つである遺伝的プログラミング (Genetic Programming : GP) で画像処理フィルタの組み合わせを自動で最適化する手法 [Yamaguchi 12] を用いて細胞領域を分割し、細胞の特徴量を計測するシステムの構築を行った [Hiroyasu 12]。本システムでは、画像の一部を学習領域とし、その学習領域における細胞領域分割が適切に行えるように複数の画像処理フィルタを GP を用いて組み合わせた後、その組み合わせを未学習の領域に適用する。しかしながら、GP では学習領域の分割精度が向上するように学習が行われるため、未学習領域の画像特徴量が学習領域の特徴量と異なっている場合には、その分割精度は必ずしも学習領域の分割精度に比例しない。また、領域によっては分割精度が著しく低くなるといった問題点がある。この問題に対する解決法の一つとして、複数の学習領域から複数の画像処理フィルタの組み合わせを作り、各未学習領域に適した組み合わ

せを適用する方法が考えられる。この方法において、未学習領域に対して用意された複数の画像処理フィルタの組み合わせのうち、どの組み合わせを適用すればよいかを決定するのは困難である。そこで、本稿では画像のテクスチャ特徴量を用いたクラスタリングを行うことによって、学習領域と学習した木構造状フィルタを適用する未学習領域を決定する手法を提案する。また、最適なテクスチャ特徴量の組み合わせを探索するために、遺伝的アルゴリズム (Genetic Algorithm : GA) を用いた探索手法を提案する。

## 2. 角膜内皮細胞画像特徴量自動抽出システム

### 2.1 システム概要

培養された角膜内皮細胞において、細胞の密度や面積のばらつき、形状といった特徴量は細胞の状態を確認するうえで重要になる。しかし、顕微鏡によって撮影した画像から直接特徴量を計測することは困難である。そこで、我々は図 1 に示すような細胞の特徴量自動抽出システムを提案している [Hiroyasu 12]。本システムは以下に示す手順で利用する。

#### Step.1 システムへの入力

ユーザーは、対象となる細胞画像とそれをもとに作成した学習セットをシステムに入力する。学習セットの学習領域は細胞画像の一部を切り取ったものである。また、理想の分割結果はユーザーがペイントソフトなどを用いて手作業で作成する。

#### Step.2 GP を用いた木構造状フィルタの構築

システムは入力された学習セットを用いて、GP で既存の画像処理フィルタを最適に組み合わせた木構造状フィルタを構築する [Yamaguchi 12]。

#### Step.3 特徴量抽出結果の出力

システムは学習した木構造状フィルタを全体画像に適用し、領域分割を行う。その結果をもとに細胞の特徴量を計測し、ユーザーに提示する。

### 2.2 評価関数

本システムでは、GP の評価関数として、領域分割の定量的な評価で用いられる GCE(Global Consistency Error)[Tal 01] を使用する。GCE は理想の分割結果  $T$  と構築された木構造状フィ

連絡先: 関谷駿介, 同志社大学大学院生命医科学研究科, 京都府京田辺市多々羅都谷 1-3, ssekiya@mis.doshisha.ac.jp

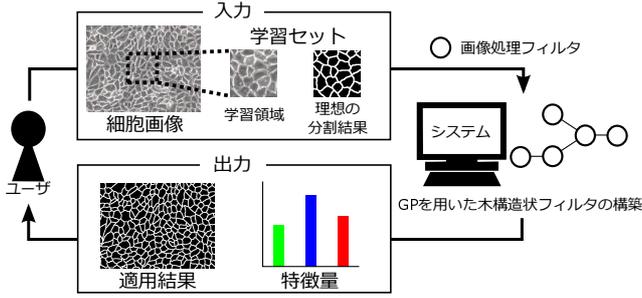


図 1: 特徴量自動抽出システム

ルタによる出力画像  $O$  上のピクセル  $p_i$  が属する細胞領域に注目し、その細胞領域を構成するピクセル数 ( $R(T, p_i), R(O, p_i)$ ) の誤差 (局所的詳細誤差)  $E$  で表すことができる。局所的詳細誤差を式 (1) に示す。この局所的詳細誤差を画像内の全てのピクセルごとに調べることで、画像全体の分割精度である GCE を評価することができる。GCE は  $N$  を全ピクセル数とすると式 (2) で定義される。

$$E(T, O, p_i) = \frac{|R(T, p_i) \setminus R(O, p_i)|}{|R(T, p_i)|} \quad (1)$$

$$GCE(T, O) = \frac{1}{N} \min \left\{ \sum_i E(T, O, p_i), \sum_i E(O, T, p_i) \right\} \quad (2)$$

### 2.3 木構造状フィルタによる領域分割の問題点

本システムの学習セットに対する有効性は既に確認されている [Hiroyasu 12]。しかしながら、構築された木構造状フィルタの未学習領域の分割精度は、領域によっては著しく低くなるという問題がある。この問題に対して、複数の学習領域から複数の木構造状フィルタを構築し、それぞれを適切な未学習領域へ適用する解決法が考えられる。細胞画像は図 2 のように、ある画像特徴量を用いると未学習領域の分割精度が学習領域の分割精度に比例する複数のクラスを形成できると仮定できる。

そこで、本稿では画像のテクスチャ特徴量に着目し、未学習領域の分割精度とテクスチャ特徴量の関係から未学習領域の分割精度が安定するような複数の学習領域を決定する手法を提案する。次章ではクラスを形成するために使用するテクスチャ特徴量及び、それを用いた学習領域の決定法について述べる。

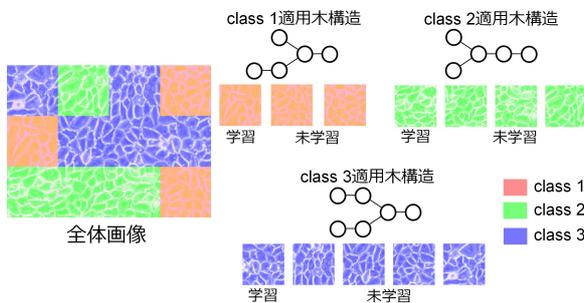


図 2: クラスタリングによる木構造状フィルタの分配

## 3. 提案手法

本手法では、未学習領域の分割精度が学習領域の分割精度に比例する複数のクラスを形成するための特徴量として、画像のテクスチャ特徴量を用いる。テクスチャとは画素値の 2 次

元的な変化パターンであり、画素分布に関する統計的な性質や空間周波数成分を分析することで表すことができる。本手法ではテクスチャ特徴量である濃度共起行列 (Spatial Gray Level Dependence Method: SGLDM) [Haralick 73], 高次局所自己相関 (High-order Local Auto-Correlation: HLAC) 特徴量 [Kurita 92], Local Binary Pattern (LBP) 特徴量 [Ojala 96], Wavelet 特徴量 [Mallat 89] から算出される合計 66 個の特徴量を用いる。また、このテクスチャ特徴量の最適な組み合わせを GA を用いて探索する。その後、探索した特徴量の組み合わせを用いて階層型クラスタリングを行い、各クラスターの重心に最も距離が近い領域を学習領域とする。

### 3.1 GA を用いたテクスチャ特徴量の組み合わせ最適化

学習領域に対するテクスチャ特徴量空間の距離が近いほど、未学習領域が高い分割精度を有するためには、図 3 の右図のように、未学習領域分割精度が学習領域に対する特徴量の距離と反比例するような特徴量の組み合わせを用いてクラスタリングを行う必要がある。本手法では、GA を用いて前節で述べたテクスチャ特徴量合計 66 個の組み合わせを探索する。その際、未学習領域の数を  $n$  とすると、分割精度で降順ソートした未学習領域分割精度  $A_1, A_2, \dots, A_n$  と、学習領域に対するテクスチャ特徴量空間のユークリッド距離で昇順ソートした未学習領域分割精度  $a_1, a_2, \dots, a_n$  の距離 (式 (3)) を GA の評価関数として用いる。

$$fitness = \sqrt{\sum_{i=1}^n (A_i - a_i)^2} \quad (3)$$

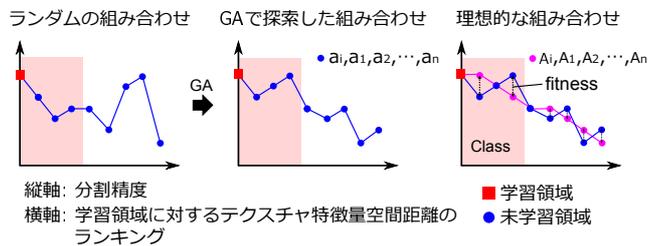


図 3: GA によるテクスチャ特徴量の組み合わせの探索

### 3.2 テクスチャ特徴量を用いた階層型クラスタリングによる学習領域の決定法

提案手法では、学習領域候補を任意のサイズで全体画像を区切った各領域とする。次に、前節の手法で探索したテクスチャ特徴量を用いて階層型クラスタリングを行い、学習領域候補を複数のクラスターに分類する。目的のクラスター数までクラスタリングが完了すると、各クラスターの重心に最も距離が近い学習領域候補を学習領域として取得する。また、クラスター同士の距離計算方法には Ward 法を用いる。Ward 法では、クラスターを  $C$ 、クラスターを形成するデータを  $x$ 、クラスターの重心を  $c$ 、距離を  $D$ 、各クラスターにおける各データとの偏差平方和を  $E$  (式 (4)) とすると、クラスター間の距離は式 (5) によって計算される。このとき、距離  $D$  はユークリッド距離を用いる。

$$E(C_i) = \sum_{x \in C_i} (D(x, c_i))^2 \quad (4)$$

$$D(C_1, C_2) = E(C_1 \cup C_2) - E(C_1) - E(C_2) \quad (5)$$

## 4. 実験

### 4.1 実験概要

本実験では、実験画像としてカニクイザルから採取した角膜内皮細胞を培養し、位相差顕微鏡によって 200 倍で撮影した図 4 のような  $400 \times 300$ pixel の細胞画像を 4 枚利用する。以下に実験手順を述べる。

#### Step.1 学習・未学習領域の分割精度の取得

図 4 からそれぞれ 3 枚の学習領域を  $100 \times 100$ pixel のサイズで点線で囲まれたようなランダムな位置 (Random 1~3) から切り取り、それぞれに対して 2.1 の Step.2 で述べた手法で木構造状フィルタを構築後、未学習領域に  $100 \times 100$ pixel 毎に適用する。

#### Step.2 GA によるテクスチャ特徴量の組み合わせの探索

Step.1 を 7 試行行った後、各未学習領域分割精度の中央値を用いて GA の学習に用いる図 3 の右図のような分割精度で降順ソートしたデータを作成する。各細胞画像につき 3 個のソートデータがあるので、式 (3) の総和が最少となるような特徴量の組み合わせ (Proposal 1) を GA を用いて探索する。また、全細胞画像合計 12 個のソートデータを用いた特徴量の組み合わせ (Proposal 2) の探索も行う。

#### Step.3 学習領域の取得

図 4 の画像を  $100 \times 100$ pixel のサイズで区切り、学習領域候補とする。そして、Step.2 で探索した Proposal 1 と Proposal 2 のそれぞれ特徴量の組み合わせを用いて 3 クラスまで階層型クラスタリングを行い、各クラスターの重心に最も距離が近い候補を学習領域とする。

#### Step.4 提案手法の有効性の検証

Step.3 で取得した各学習領域に対して構築した木構造状フィルタを同クラスターの未学習領域に適用する。この時の未学習領域の平均分割精度の最大値を初めにランダムに決定した各学習領域 (Random 1~3) の場合及び、66 個全ての特徴量でクラスタリングした場合 (All) と比較する。

なお、ここで用いる分割精度は 1.0 から GCE を引いた値とし、学習の際に用いる理想の分割結果は、角膜再生医療の研究者がペイントソフトを用いて学習領域における細胞境界を白、細胞領域を黒に手作業でマーキングしたものとする。実験に用いた GP・GA のパラメータ及び、使用した画像処理フィルタを表 1~表 3 に示す。

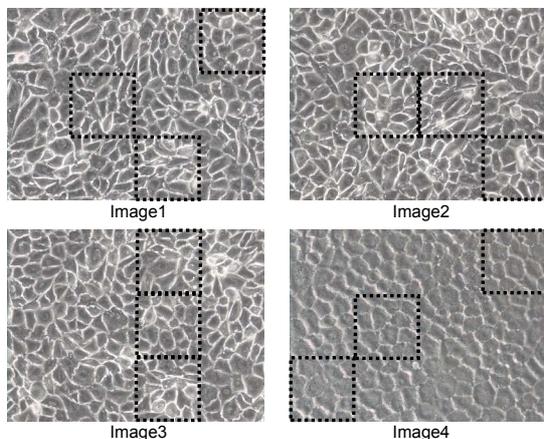


図 4: 実験画像

表 1: GP・GA のパラメータ

パラメータ	GP	GA
個体数	300	50
世代数	300	100
交叉率	1.0	1.0
突然変異率	0.01	0.01
選択方法	トーナメント	トーナメント
トーナメントサイズ	4	2
ペナルティ	$1.0 \times 10^{-3}$	$1.0 \times 10^{-3}$
試行回数	7	10
遺伝子数		66
交叉方法		二点交叉

表 2: 1 入力フィルタ

number	フィルタ名
f1	最小値
f2	最大値
f3	収縮
f4	膨張
f5	平均値
f6	中央値
f7	Gaussian
f8	2 値化 (大津)
f9	反転
f10	孤立領域 (小)
f11	孤立領域 (大)
f12	Watershed
f13	Watershed(距離変換)
f14	Watershed(2 値)
f15	Watershed(距離変換・2 値)

表 3: 2 入力フィルタ

number	フィルタ名
F1	論理和
F2	論理積
F3	代数和
F4	代数積
F5	限界和
F6	限界積

### 4.2 実験結果

Step.2 において、10 試行中最も評価値が高かったテクスチャ特徴量の組み合わせを表 4 に示す。また、その時の GA による評価値と、全てのテクスチャ特徴量を用いたときの評価値を表 5 に示す。表 4 と表 5 より、SGLDM に関する特徴量が多く選択されること及び、GA の探索が上手く進んだことが確認できた。

Step.3 において、表 4 の特徴量の組み合わせによってクラスタリングした実験画像をクラスター毎に色分けして図 5 と図 6 に示す。この結果より、各クラスターに平等な数の領域が割り振られていないことが確認できる。また、学習領域図 5、図 6 の点線で示した領域となった。

Step.4 の結果を表 6 に示す。この結果より、Proposal 1,2 は、Random 1~3 及び All の場合との分割精度の差が比較的大きいため、学習した木構造状フィルタを適切な未学習領域に配分できる学習領域を決定できたといえる。

表 4: GA による特徴量探索結果

画像	特徴量
Image 1	エントロピー (SGLDM 45°), 慣性 (SGLDM 0°), 尖度 (LBP), 1 レベル HL 成分 (Wavelet)
Image 2	エネルギー (SGLDM 35°), 相関 (SGLDM 0°, 135°), 歪度 (LBP)
Image 3	エネルギー (SGLDM 45°), 相関 (SGLDM 45°), 尖度 (LBP)
Image 4	HLAC33, HLAC34, 平均 (LBP)
全画像	エネルギー (SGLDM 45°, 90°), エントロピー (SGLDM 90°, 135°), 相関 (SGLDM 90°), 2 レベル HH 成分 (Wavelet), 2 レベル HL 成分 (Wavelet), 3 レベル HH 成分 (Wavelet)

表 5: GA による特徴量探索時の評価値

画像	Proposal 1	Proposal 2	All
Image 1	0.98	1.16	1.09
Image 2	1.14	1.32	1.80
Image 3	0.85	1.16	1.26
Image 4	1.05	1.05	1.58

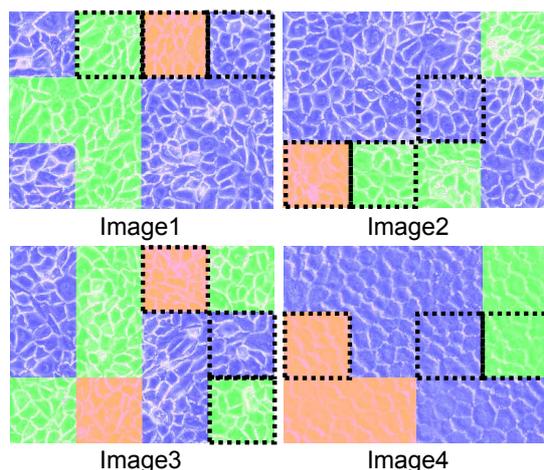


図 5: Proposal 1 の組み合わせによるクラスタリング結果

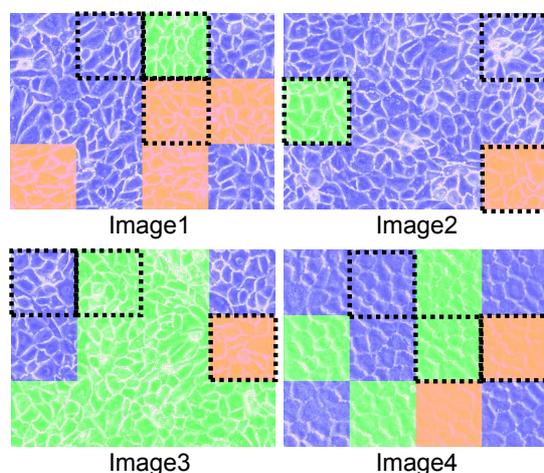


図 6: Proposal 2 の組み合わせによるクラスタリング結果

表 6: 未学習領域分割精度の比較

画像	Proposal 1	Proposal 2	Random 1	Random 2	Random 3	All
Image 1	0.74	0.74	0.56	0.71	0.69	0.68
Image 2	0.65	0.64	0.65	0.65	0.64	0.64
Image 3	0.62	0.64	0.58	0.61	0.63	0.64
Image 4	0.61	0.59	0.57	0.55	0.55	0.56

## 5. 考察

前章より、提案手法の有効性を確かめることが出来た。しかしながら、提案手法が劣っている場合があること、また、図 5、図 6 に示すようにクラスタを構成する領域数に偏りが生じていることなどが問題点として挙げられる。この結果の原因として、特徴量の距離を考慮した GA の評価関数を設定しなかったことが考えられる。図 7 に横軸を学習領域に対する特徴量空間の距離で表した、Image 3 の GA 探索時における結果を示す。提案手法では理想的な降順ソートとの分割精度の距離のみで評価を行っているため、選択された特徴量の組み合わせでクラスタリングを行うと図 7 のように 1 クラスタ当たりの領域数に偏りが生じてしまう。そのため、学習した木構造状フィルタを効率的に適用することができなかつたと考えられる。今後は、GA による探索時に 1 クラスタ当たりの領域数を考慮した評価方法を検討する必要がある。また、探索されたテクスチャ特徴量の組み合わせが他検体に対しても有効なのかを調査する必要がある。

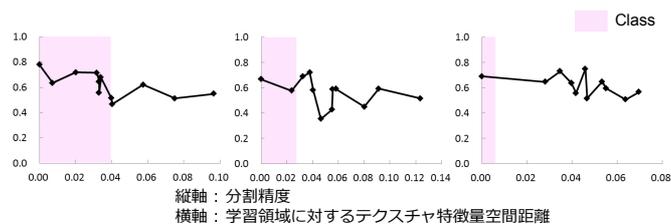


図 7: GA によって探索した特徴量空間距離によるソート結果

## 6. まとめ

本稿では、細胞領域分割のための画像処理 GP における未学習領域分割精度を安定させることを目的とし、学習・未学習領域分割精度と画像のテクスチャ特徴量の関係を用いて学習領域を決定する手法を提案した。本手法では、あらかじめ取得した学習・未学習領域分割精度を基に、学習した木構造状フィルタを適切な未学習領域に適用できるようにクラスタを形成するためのテクスチャ特徴量の組み合わせを GA を用いて探索する。実験として、提案手法で得られた学習領域とランダムで選択した学習領域及び、全テクスチャ特徴量によるクラスタリングで決定された学習領域における未学習領域分割精度の比較を行った。結果、提案手法の分割精度が大きく上回る場合もあり、その有効性を確認することができた。今後は、テクスチャ特徴量の組み合わせ探索時に用いる GA の評価方法の改善と探索した組み合わせの他検体への有効性を調査を行う必要がある。

## 参考文献

- [Koizumi 09] N. Koizumi: Cultivated corneal endothelial cell sheet transplantation in a primate model, Journal of Japanese Ophthalmological Society, vol. 113, no. 11, pp. 1050-1059, 2009.
- [Yamaguchi 12] H. Yamaguchi, et al: Comparison study of controlling bloot model of GP in constructing filter for cell image segmentation problems, in Evolutionary Computation (CEC), 2012 IEEE Congress on, 2012, pp. 1-8.
- [Hiroyasu 12] T. Hiroyasu, et al: Algorithms for automatic extraction of feature values of corneal endothelial cells using genetic programming, in Soft Computing and Intelligent Systems (SCIS) and 13th International Symposium on Advanced Intelligent Systems (ISIS), 2012 Joint 6th International Conference on, 2012, pp. 1388-1392.
- [Tal 01] D. Tal, et al: A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics, IEEE International Conference on Computer Vision, vol. 2, pp. 416-423, 2001.
- [Haralick 73] R. Haralick, et al: Textural features for image classification, Systems, Man and Cybernetics, IEEE Transactions on, vol. SMC-3, no. 6, pp. 610-621, 1973.
- [Kurita 92] T. Kurita, et al: A face recognition method using higher order local autocorrelation and multivariate analysis, in Pattern Recognition, 1992. Vol.II. Conference B: Pattern Recognition Methodology and Systems, Proceedings. 11th IAPR International Conference on, 1992, pp. 213-216.
- [Ojala 96] T. Ojala, et al: A comparative study of texture measures with classification based on featured distributions, Pattern Recognition, vol. 29, no. 1, pp. 51-59, 1996.
- [Mallat 89] S. Mallat: A theory for multiresolution signal decomposition: the wavelet representation, Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 11, no. 7, pp. 674-693, 1989.