

# 遺伝的アルゴリズムを用いた画像探索システムの研究

布 広 永 示<sup>†</sup> 片 山 啓<sup>††</sup>

<sup>†</sup> 東京情報大学 <sup>††</sup> 東京情報大学大学院

画像データを有効利用するためには、必要なデータを画像データベースの中から素早く探し出す探索機能が必要である。本研究は、東京情報大学における学術フロンティアの研究成果として蓄積されている地球観測データである MODIS 画像の中から必要な画像データを探索する画像探索システムを開発する。本システムの画像探索アルゴリズムとしては、確率的探索手法の一つである遺伝的アルゴリズム (Genetic Algorithms: GA) を用い、画像データを複数の PC に分散して処理するクラスタ環境における遺伝的アルゴリズムの適用性について評価した。

## Research of image search system that uses Genetic Algorithm

Eiji NUNOHIRO<sup>†</sup> and Kei KATAYAMA<sup>††</sup>

<sup>†</sup> Tokyo University of Information Science

<sup>††</sup> Graduate School of Tokyo University of Information Science

A search function to find out the significant data which user needs quickly from image database is necessary to use image data effectively. This research develops image search system to find the useful image data from the database of MODIS (new earth observing system in Terra satellite launched by NASA) data which set up in Tokyo University of Information Sciences as part of Academic Frontier project. For image search algorithm of this system, we use Genetic Algorithms (GA) which is one of the probabilistic search technique and evaluate the adaptability of GA for searching the image data. Furthermore, we research speedup of search time using GA in parallel distributed computing environment by cluster PC.

### 1 はじめに

画像データを有効利用するためには、必要なデータを画像データベースの中から素早く探し出す探索機能が必要である。本研究は、東京情報大学における学術フロンティアの研究成果として蓄積されている地球観測データである MODIS 画像の中から必要な画像データを探索する画像探索システムを開発する。本システムの画像探索アルゴリズムとしては、確率的探索手法の一つである遺伝的アルゴリズム (Genetic Algorithms: GA) を用い、画像データを複数の PC に分散して処理するクラスタ環境における遺伝的アルゴリズムの適用性について研究する。

### 2 画像探索システムの概要

東京情報大学では、Terra/Aqua 衛星に搭載されているセンサーの一つである MODIS の画像データを受信して、大学や研究機関に配信している。画像データを有効利用するためには、利用者にとって必要な画像の位置、領域、分布などに関する情

報を大量の画像データベースの中から素早く抽出する探索機能が必要である。本研究では、画像データの中から特定の形状や特性を持つ領域を抽出する画像探索システムを開発する。本画像検索システムにおいては、探索処理の向上化のために、画像を分散して並列に処理する方法を考える。ここで、画像の分散方法は、次の 2 つに分けられる。①一つの画像を複数の領域に分割して、分割した部分画像毎に探索処理を実施する。②大量の画像を幾つかの画像群にグループ化して、グループ毎に探索処理を実施する。本研究では、画像の分散方法として、1つの画像を複数の領域に分割して、各領域に対して同時に探索処理を実施し、各領域のどれかで探索画像が抽出された時点で探索処理が終了する並列探索モデルを考える。また、画像探索アルゴリズムとしては、確率的探索手法の一つである遺伝的アルゴリズム (Genetic Algorithms: GA) を用いる。画像探索システムの概要を図 1 に示す。

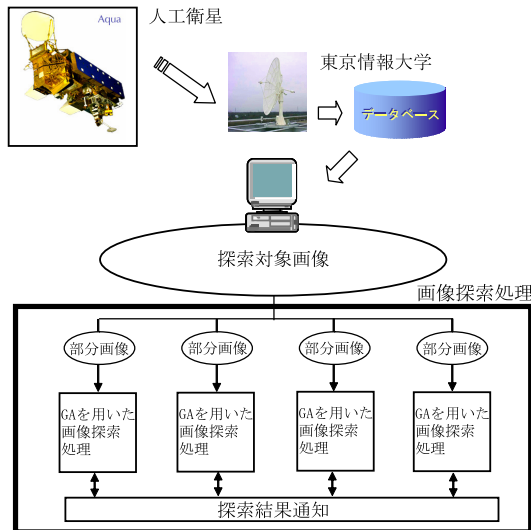


図 1 画像探索システムの概要

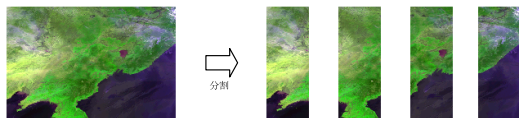


図 2 画像分割のイメージ

### 3 画像分割方法

探索対象の画像は部分画像に分割されて、部分画像単位に探索処理が実施される。ここでは、探索対象の画像を部分画像に分割する方法について記述する。画像の分割イメージを図 2 に示す。

画像を分割し、分割された部分画像に対して探索処理を並列に実施することで探索処理時間の短縮を図る。しかし、図 3 に示すように、分割した画像の境界に探索画像が存在している場合には、目的とする画像を探索が出来ないといった問題が発生する。このため、画像を単純に分割するのではなく、探索画像の領域サイズを考慮した処理が必要となる。

そこで、本システムでは、探索画像のサイズに合わせて部分画像に拡張領域（オーバラップ領域）を付加することで対応した。例えば、図 4 に示すように、部分画像の縦・横のサイズが  $x_1, y_1$  とし、探索画像の縦・横のサイズが  $x_2, y_2$  とした場合、部分画像の分割面のサイズを縦方向に  $x_2$ 、横方向に  $y_2$  だけ拡張して分割する。しかし、この方法では探索画像より大きい領域が分割された場合の探

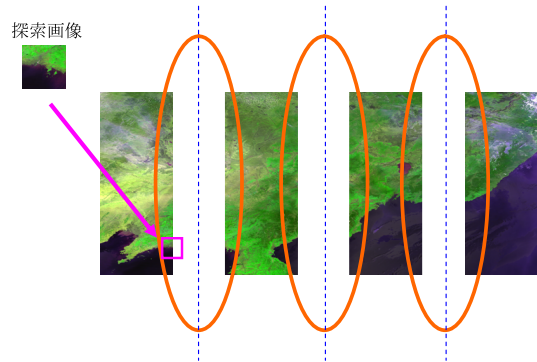


図 3 画像の分割パターン

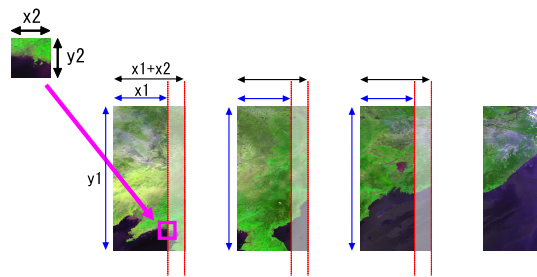


図 4 画像分割のパターン

索が出来ない。拡張領域を動的に変化させるなどの対応が必要であるが、この対応については別途検討することにする。

### 4 遺伝的アルゴリズムの処理内容

本研究では、探索画像の探索対象画像における位置、拡大倍率、回転角度を変化させた領域を生成し、その領域の中で画素の階調値のマッチング率が最適な領域を抽出することを考える。したがって、画像探索問題を、画素の階調値のマッチング率を最大化する領域の中心座標  $(x, y)$ 、拡大倍率  $rate$ 、回転角度  $angle$  の 4 つのパラメータを決定する問題として GA を適応する。そして、4 つのパラメータを染色体とする遺伝子を生成し、各個体のマッチング率を適応度とする。

#### 4.1 遺伝的アルゴリズムの概要

GA とは生物の進化機構を模倣した最適化アルゴリズムであり、解候補集合を世代交代させることで漸近的に近似解に到達する。GA では、探索空間上の探索点を生物個体とみなす。固体の母集団 (population) に対して、選択 (selection)、交叉 (crossover)、突然変異 (mutation) という遺伝

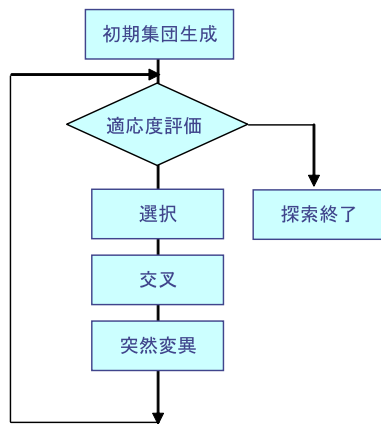


図 5 GA の基本的処理の流れ

x	y	rate	angle
---	---	------	-------

図 6 遺伝子の染色体

的操作 (genetic operator) を繰り返し適応することにより、環境への適応度 (fitness) が向上し、最適解に到達することが期待される。GA の基本的な処理の流れを図 5 に示す。

#### 4.2 染色体上の表現方法

1 つの遺伝子を表現する染色体を探索画像に対する探索領域の中心座標 (x, y), 拡大倍率 rate, 回転角度 angle とする。染色体の内容を図 6 に示す。また、探索画像に対する探索領域の関係を 7 に示す。

#### 4.3 選択、交叉方法と突然変異作用法

選択方法、交叉方法、突然変異作用法としては、色々な方法が提案され、探索の効率は、個体へのコーディング、交叉・突然変異の設計、世代交代手法 (複製選択と生存選択) に依存する。本研究では、次のような方法で評価を開始し、その評価結果によって、別の方法も比較検討することとした。

##### (a) 選択方法

選択方法として、集団の中で最も適応度の高い個体をそのまま次世代に残すエリート保存戦略を用いて評価を実施する。

##### (b) 交叉方法

交叉方法として、染色体の切断箇所をランダムに 1 カ所指定し、その箇所での親の遺伝子を交叉させる 1 点交叉法を実施する。

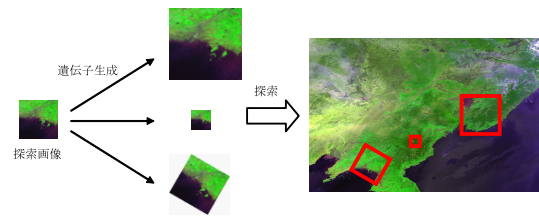


図 7 画像分割のパターン

##### (c) 突然変異作用素

突然変異作用素として、ランダムに選ばれた 2 点間の遺伝子順序を逆に並べ替える方法 (逆位) を実施する (突然変異率: 0.05)。

#### 4.4 遺伝的アルゴリズムの手順

GA とは、初期集団を発生させた後、適応度の評価、選択 (淘汰)、交叉、突然変異を繰り返して最適解を得ようとするものである。本研究における GA の手順は次のとおりである。

- ① 分割点群の生成 (固体群の生成)  
0,1 で構成されるランダムな染色体 (chromosome) を持つ個体を n 個生成する。
- ② 適合度 (fitness) の評価  
適合度関数 (fitness function) により適用度を評価する。
- ③ 終了判定 (条件を満足しない場合は、②の繰り返し)
- ④ 選択 (selection)  
個体の適合度を考慮のうえ、個体のペアを確率的に n 組生成する。
- ⑤ 交叉 (crossover)  
設定された交叉方法に従って、各ペアを親として、新しい (子の) 個体を生成する。
- ⑥ 突然変異 (mutation)  
設定された突然変異確率に従って、各子の個体ごとに、染色体上の遺伝子の値を他の値に置き換える。
- ⑦ ②に戻る。

#### 4.5 適応度の評価方法

適応度の評価は、探索画像に対する探索対象画像の探索領域の中心座標  $(x, y)$ 、拡大倍率  $rate$ 、回転角度  $angle$  における画素の階調値のマッチング率を評価することである。各画素は RGB の各色 0~255 の 256 階調値 (各 8 ビットで計 24 ビット) で表現しており、適応度は各色の階調値の差の合計を評価する。探索領域における探索画像の画素を  $img1$ 、探索対象画像の画素を  $img2$ 、探索領域の画素数を  $x*y$  とした時の適応度関数を式 (1) に示す。

$$fitness = 1.0 - \left( \sum_i^y \sum_j^x \sum_k^3 |img1_{ijk} - img2_{ijk}| \right) / (y * x * 3 * 255)$$

ここで、 $| \quad |$  は絶対値を表す。

この適応度関数の値を評価して、値が一定の基準 (最大適応度) を超える領域をマッチング率が最適な領域として抽出する。

#### 5 システム構成

画像探索システムの構成を図 8 に示す。データ管理サーバでは、画像情報管理、画像分割・配信、GA パラメータの設定・配信を行う。PC クラスタは、部分画像と GA パラメータを受信して画像探索処理を行う。また、画像探索システムの実装言語としては、異機種混在型のクラスタ環境におけるアプリケーションのポータビリティを考慮して、Java 言語を採用し、データの通信や実行制御を行う通信手段として Java RMI (Java Remote Method Invocation) を利用した。

#### 6 システムの評価結果

表 1 に示す条件で、性能の比較実験を行った。性能評価測定方法は、原画像から検索画像を発見した時点の経過時間を取得し、これを 1 回の試行と数えてる。この操作を 100 回行い、平均値を取ったものを性能とする。利用したコンピュータ環境は、HewlettPackard 社 nx6120/CT, CPU: Pentium M 740 (1.73GHz), 主記憶: PC2700 DDRSDRAM 512MB である。

表 2 に性能評価結果、図 9 に性能向上比、世代数比を示す。この結果に示すように、クライアン

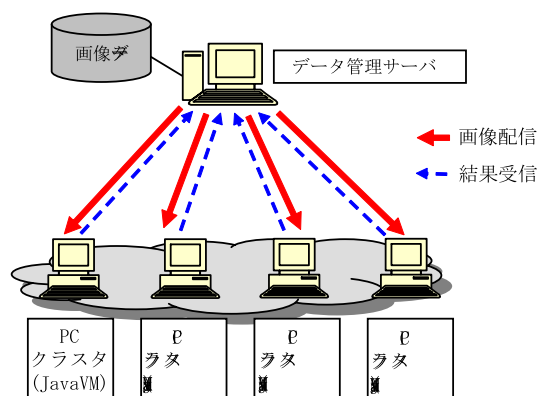


図 8 画像検索システムの概要

表 1 実験における環境条件

検索ファイル	原画像ファイルから任意に抽出したデータ (固定)
原画像ファイル	MODIS 画像 (TIFF データ)
GA における個体総数	200
GA における世代数	500
GA における淘汰法	各世代で算出された適応度の比較
GA における交叉法	一点交叉
GA における交叉率	0.5
GA における突然変異率	0.05
GA における最大適応度	0.92
GA におけるエリート保存戦略	保存戦略を行う

トの台数に比例して、探索処理に要する時間、世代数が改善されていることが確認された。ただし、クライアント数が増えると、性能向上比が小さくなる傾向にあるので、この点については今後改善する必要がある。

#### 7 おわりに

本件では、衛星画像データの中から特定の画像を抽出する画像探索を、画像の領域を分割し、遺伝的アルゴリズムを用いて各領域を並列に処理することで実現した。今後の研究テーマとしては、複雑な形状や特性を持つ領域の探索方法に対する並列分散型の遺伝的アルゴリズム (Parallel DGA: PDGA) の適用性の評価や時系列における地形の

表 2 性能評価結果

クライアント数	経過時間	性能向上比 (時間)	平均世代数	性能向上比 (世代数)
1 Client	11577 ms	—	157.8	—
2 Client	4067 ms	2.8	55.6	2.8
4 Client	2315 ms	5.0	32.6	4.8
8 Client	1674 ms	6.9	20.8	7.6

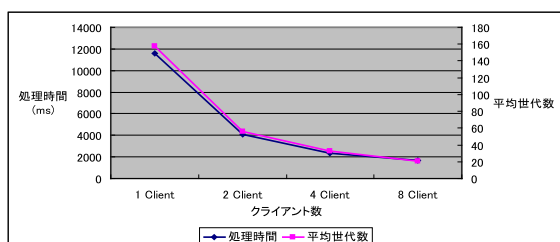


図 9 クライアント数と性能向上比，平均世代数の比較

形状変化に対する遺伝的アルゴリズムの適応性の評価などを考えている。

謝辞 本研究は，私立大学学術研究高度化推進事業 東京情報大学学術フロンティア「東アジアにおける陸圏・水圏を統合した環境情報システムの研究」の補助を受けた。

### 参考文献

- 1) 北本朝展，高木幹雄：類似画像検索への応用を目的とした階層化属性付きグラフマッチングの高速化，画像の認識・理解シンポジウム (MIRU'96)(1996),vol.2,pp331-336
- 2) 北本朝展，高木幹雄：類似画像検索システム構築のフレームワークとしての階層モデル，電子情報通信学会技術報告 (1997),PRMU97-58,pp25-32.
- 3) 北本朝展，高木幹雄：進化的計算論に基づく対話的な画像散策法，第 4 回知能情報メディアシンポジウム (1998),pp173-180
- 4) 廣安知之，三木光範，佐野正樹，谷村勇輔：2 個体分散遺伝的アルゴリズム，情報処理学会第 5 回数理モデル化と問題解決シンポジウム論文集 (2000), pp.49-56.