

ボトムアップ手法によるグラフ分割を用いた染色体の検出

高見 燎世^{1,a)} 槇原 靖¹ 八木 康史¹ 川端 絵美² 衣笠 泰葉² 田代 聡²

概要：本研究では、染色体画像全体からボトムアップ手法を用いて染色体を検出する手法を提案する。まず、スケールスペース上での極値探索により特徴量抽出を行う SIFT (Scale-Invariant Feature Transform) を用いて染色体の部位であるセントロメアとテロメアを検出する。次に、検出された部位間を接続する DNA に対するスコア (信頼度) を、青成分の値を用いて行うスコアを計算する。最後にセントロメアやテロメアの特徴点としてのスコアと DNA のスコアを用いて、染色体画像全体に対するグラフを構築して、グラフ分割により個別の染色体への分割を実現する。実験では、セントロメア・テロメアの検出の再現率と精度の評価、及び染色体全体としての再現率と精度を評価した。その結果、セントロメアでは 82.3%、テロメアでは 95.6%の精度を、染色体の検出では 93.3%の精度を得た。

キーワード：染色体, DNA, セントロメア, テロメア, 細胞画像処理, ボトムアップ手法, グラフ分割, SIFT

1. はじめに

顕微鏡技術や蛍光技術の発展により、細胞画像処理による生命現象の解明が日々着々と進んでおり、脳疾患に関わる細胞や癌化の原因細胞の特定等の試みがなされている [1], [2], [3]。しかし、その裏側では、実験担当者による目視によるデータ解析を伴うことも多く、研究者のエフォートを単調作業（細胞数のカウントや細胞のサイズの計測、細胞追跡のための画像間の細胞対応付け）に割かざるを得ないことも多々ある。更には、人手による目視確認では系統的・定量的な解析をするには限度が有り、無数の細胞を対象とした追跡は、人手では事実上不可能であり、結果として、生命現象の解明の障壁となってしまっている。そのため、対象に応じた適切な画像処理手法 [4], [5] を開発・適用することで、自動かつ定量的に細胞の特性や挙動を評価することが可能となり、生命科学分野の研究者の単調作業へのエフォートの低減や、従来の手動処理では解析しきれなかった現象の解明へと繋がるのが期待される。

そのような現象の一例として、放射線による染色体への解析が挙げられる。放射線の高線量照射により染色体に異常をきたすことはよく知られているが、福島第一原発での事故による周辺住民の低線量の放射線被曝や CT スキャン

による白血病や悪性腫瘍の発症 [6] などのように、低線量の放射線による染色体異常の現象については、十分に解明されているとは言えない。

染色体は細胞の細胞核内に 46 本存在し、それら染色体には DNA の交差する部位であるセントロメアが 1 つと染色体の末端部にある部位であるテロメアが 4 つ、それら部位を DNA がつながり合わせている形状をしている。放射線照射によって DNA が切断され修復しようと DNA と結合を行い、この際誤った DNA と結合をすることで染色体形状に異常が起き人体に悪影響を及ぼすことがわかっている [7]。図 1 のように通常の染色体は 1 つのセントロメアと 4 つのテロメアに接続しているが、セントロメアが 2, 3 個接続した形をしている dicentric, Tricentric やセントロメアが一つのみの Ring や Telomere 同士が繋がったペアである fragment のような染色体異常が起きる。染色体異常の解析のために、既に染色体蛍光画像を対象とした自動解析ソフト Metafer が開発されているが、染色体の検出の失敗や、染色体異常の種類の認識別が一定程度含まれており、低線量放射線による染色体異常を定量的に解析するには不十分である [8]。

染色体においてセントロメアとテロメアを関節、DNA をリンクとみなすと、ある種、関節物体の検出や構造推定の問題とみなすことができる。そのため、同じ関節物体である人体の姿勢推定と技術的な類似点があると考えられる。人物姿勢推定とは画像や動画像に対して人間の関節（肘、膝など）の場所の特定、及びポーズの分類を行うものとして定

¹ 大阪大学
Osaka University

² 広島大学
Hiroshima University

^{a)} takami@am.sanken.osaka-u.ac.jp

染色体異常の種類

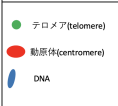
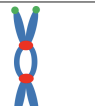
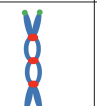
	正常染色体	二動原体染色体 (dicentric chromosome)	三動原体染色体 (tricentric chromosome)	環状染色体 (ring chromosome)	fragment
					
オリジナル画像 (赤蛍光：動原体； 緑蛍光：テロメア； 青蛍光：DNA)					
自動解析ソフト Metafer で認識された染色体 (赤蛍光：動原体； 緑蛍光：テロメア； 白黒：DNA)					

図 1 染色体異常による形状変化

義される。近年人物姿勢推定分野は CNN などのディープラーニング技術の進歩によって大きな転換を迎えている。人物姿勢推定にはトップダウン手法とボトムアップ手法という二つの手法が存在する。

トップダウン手法は人物の検知を最初に行い、次にそれぞれの人物について姿勢推定を行う手法である。He らの Mask R-CNN[9] や Iqbal らの整数線型計画法 [10], Yilun らの Cascaded Pyramid Network[11] は代表的なトップダウン手法のアルゴリズムである。トップダウン手法は各人物に対する人物検出と姿勢の推定を別のステップで行うため、人物検出を誤検出した場合姿勢推定を行うことが不可能になる。また人数分姿勢推定を実行するので人数に比例して計算量が増加する。これは Wei らの CNN による個人の姿勢推定 [12] を複数人に拡張して発展したためである。

ボトムアップ手法は画像中の人間の関節（キーポイント）をすべて検出し、それぞれの人物に対してキーポイントのマッチングを行い接続する手法である。Pishchulin らの Deepcut[13] やそれを改善した Insafutdinov らの Deepercut[14] や Cao らの OpenPose[15], [16] や Insafutdinov らの Arttrack[17] のような様々なボトムアップ手法が発展し高い精度を出している。ボトムアップ手法では人物検出→姿勢推定というステップを踏まない。そのため人物検出を失敗することによって姿勢推定が不可能になる問題が起こらない、さらに物体同士の重なりに対しても画像全体として検出できる可能性がある。さらに画像全体に一度に計算を行うため計算量も抑えられうる。そのためトップダウン手法の問題に対して比較的ロバストなアルゴリズムを開発することが可能である。

トップダウン手法による染色体の検出は下東ら [18] によって行われている。染色体画像に対して Voronoi 図を生成することで染色体ごとに切り分けを行い検出し、CNN を用いて分類を行っている。この手法では、染色体の検出において染色体が重なったり、正しく切り分けることができずに検出した画像に対して分類することが不可能になっ

ている。

以上を踏まえ、本研究ではボトムアップ手法を用いて、セントロメア、テロメアの部位を検出し染色体を検出する手法を提案する。また LoG による極大・極小の判定や色相を用いて誤った接続を抑制する処理も実装している。

本論文の構成を以下に示す。まず 2 章で本研究の関連研究を紹介する。次に 3 章で提案手法について詳細に述べている。4 章で提案手法の検証と考察を行っている。最後の 5 章で本研究のまとめと今後の課題を述べる。

2. 関連研究

2.1 細胞画像処理

顕微鏡技術や蛍光技術の発展によって、膨大な量の細胞や分子の画像の撮影が可能になっている。それに伴って画像処理技術を用いて、人の判断基準に頼らない画像解析の発展が期待されている。

細胞画像処理では、対象物の検出に関する研究が多数なされている。細胞内画像処理はまだ始まったばかりの分野であり、医用画像処理などに比べて最適なアルゴリズムや開発がなされていないため、この分野では目視による観測に依存しているのが現状である。しかし、細胞内画像処理では対象が無数に存在することが多々見られ目視による検出は事実上不可能であったり精度に大きく影響する。そのため検出は細胞画像処理において重要な基本プロセスであり、サポートベクタマシンを用いた脂肪滴の検出 [19] や面積、形状、濃度情報を用いた有核赤血球の自動検出 [20] など様々なアルゴリズムが提案されている。

細胞内画像中の物体の動き方や移動速度も、解析においては重要な情報となる。したがって細胞画像中の物体を自動で追跡するプロセスも検出を応用したプロセスとして細胞画像処理による生命科学分野において重要な役割を果たしている。細胞画像処理における自動追跡としてサポートベクター回帰を用いた追跡法 [21], [22], [23] や SIFT 特徴をもとにベイズの事後確率を用いた追跡法 [24] などが提案されている。

また検出を応用し、細胞内の物質を自動分類し目的の物質を識別することも細胞の解析において重要な役割を果たす細胞画像処理手法である。赤血球形態異常の自動分類 [25] や標本内情報を用いた適応的好中球分類 [26] など目的に応じたアルゴリズムが提案されている。

以上のように細胞画像処理は目的に応じて様々なアルゴリズムが開発され始めており、様々な手法において検出が基本のプロセスであることが分かる。

2.2 人物姿勢推定におけるボトムアップ手法

1 章で述べた人物姿勢推定のボトムアップ手法について本節で詳しく説明を行う。

一つ目にボトムアップ手法のアルゴリズムで代表的なも

のに pishchulin らの手法 [13] が挙げられる。このアルゴリズムは、まず姿勢推定のためにキーポイント（肘や膝、手首など身体の部位の中心点）を検出し、その後にキーポイントをマッチングさせるというものである。まず、Adapted Fast R-CNN (AFR-CNN) [27] や Dense-CNN[28], [29] を用いてキーポイントの検出を行う。それらキーポイント同士のすべての接続を考え、次に各人物のキーポイントをマッチングさせ、組み合わせ全てから各人物の組み合わせのみを残し、最後にキーポイントの部位の代表部分のみを残すというグラフ分割を行うことで、姿勢推定を行うというものである。これによりトップダウン手法の問題である、人物検出の失敗により、姿勢推定が不可能になる問題を解決している。

また Cap らによる OpenPose[15], [16] というボトムアップ手法も開発されている。OpenPose は部位検出の信頼値である Part Confidence Map, キーポイントとキーポイントの接続の確からしさ Part Affinity Fields をもとにグラフ分割を行うものである。まず、入力画像を VGG-19[28] に入力し、特徴量を抽出する。次に、キーポイントの中心を平均位置とするガウス分布によるヒートマップを教師データとして、Part Confidence Map を学習する。更に、キーポイント同士を接続する線分から一定距離内となる領域（矩形領域）内のピクセルに対して設定するリンクの向きを表すベクトル場を教師データとして Part Affinity Fields を学習する。最後にそれら Part Confidence Map と Part Affinity Fields の合計スコアを最大化するようにグラフ分割を行うことで、姿勢推定を行う。OpenPose はトップダウン手法の問題点を解決しながら Deepcut[13] より計算量を抑えて姿勢推定を行うことを可能としている。

3. 提案手法

蛍光染色された染色体画像から、セントロメアとテロメアの検出、検出点間の DNA スコアの計算、DNA スコアをもとに作成されたグラフ分割による染色体の検出を行う。

3.1 概要

染色体の検出の手順（図 2）は以下の通りである。

1. セントロメア/テロメアの検出

入力画像に対して R, G, B の各成分画像に分割したのち、R 成分画像と G 成分画像に対して SIFT[30] を用いた特徴点検出を行うことでセントロメアとテロメアの検出を行う

2. DNA スコアの計算

検出された特徴点の中で一定距離内にあるセントロメア-テロメア間、テロメア-テロメア間、セントロメア-セントロメア間を結んだ線分から一定距離内の領域（矩形領域）を作成し、その領域内の B 成分の値の平均値を DNA スコア

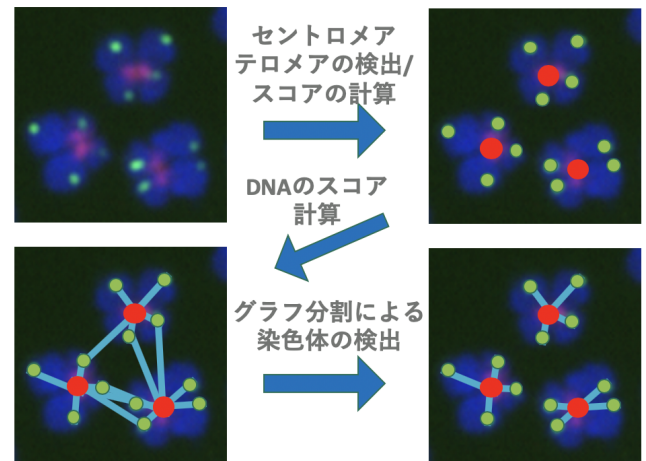


図 2 提案手法の概要（赤丸：セントロメア，緑丸：テロメア，水色線：染色体）

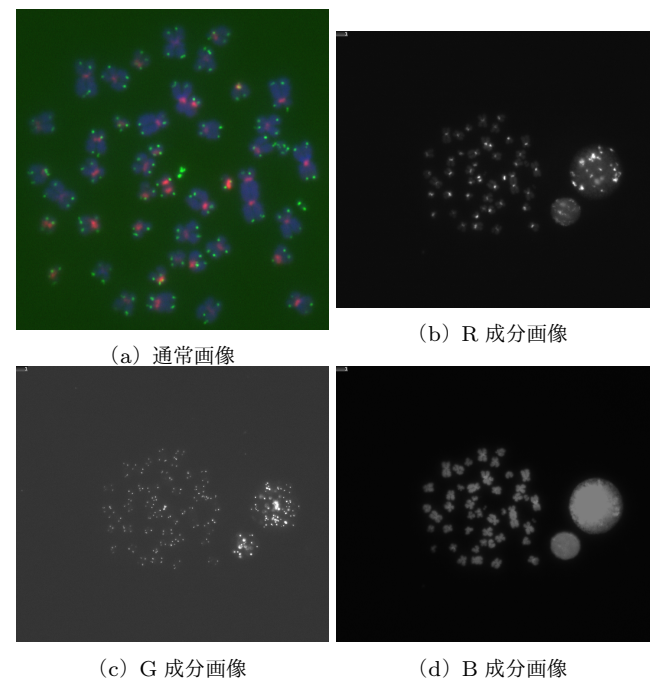


図 3 入力画像の RGB 分割

として取得する。

3. グラフ分割による染色体の検出

染色体を検出するために、スコアを全体で最大に近い結果を得るように 1 つのセントロメアに対して最大 4 つのテロメアを接続する。

3.2 セントロメア/テロメアの検出

3.2.1 RGB の分割

入力画像に対して R, G, B の各成分画像に分割を行う（図 3）。R 成分画像をセントロメア検出、G 成分画像をテロメア検出、B 成分画像を DNA スコアの取得のために用いる。

3.2.2 SIFT による特徴点検出

セントロメア及びテロメアの検出のために R 成分画像と G 成分画像を入力として SIFT を用いた特徴点検出を行う。

SIFT (Scale-Invariant Feature Transform) とは、David Lowe によって開発された特徴点検出および特徴量記述を行う画像処理アルゴリズムである [30]。SIFT の特徴点検出では DoG フィルタを用いてスケールスペースにおいて極値探索を行い、特徴点の位置とスケールを取得する、スケールスペースにおける極値探索では DoG (Difference of Gaussian) フィルタを用いて行う。DoG フィルタは LoG (Laplacian of Gaussian) フィルタを近似したフィルタである。まず LoG フィルタについて説明する。

LoG フィルタはラプラシアンフィルタとガウシアンフィルタを組み合わせたフィルタである。ガウシアンフィルタは画像の平滑化を行うことでノイズを低減することができる。ラプラシアンフィルタは 2 次微分を行うことでエッジを検出することができる。フィルタの形から円形やコーナーの特徴点を検出することが可能である。ガウシアンフィルタの定義を式 (1)、LoG フィルタの定義を式 (2) で表す。

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (1)$$

$$L(x, y, \sigma) = \frac{x^2 + y^2 - 2\sigma^2}{2\pi\sigma^6} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (2)$$

LoG は計算コストが非常に高いため DoG フィルタを用いる。DoG フィルタは異なる σ を用いてガウシアンフィルタをかけた画像の差分によって得られる。この結果は LoG(式 (2)) に近似する。DoG フィルタの定義を式 (3) で表す。 x, y, σ はそれぞれ特徴点の x 座標, y 座標, スケールである。

$$\frac{G(x, y, k\sigma) - G(x, y, \sigma)}{k\sigma^2 - \sigma^2} \simeq \nabla^2 G(x, y, \sigma) \quad (3)$$

こうして DoG を得ることができたらスケールスペースにおける極値探索を行う。1 ピクセルに対して近隣の 8 ピクセルに加えて、上下の σ に対する近隣のそれぞれ 9 ピクセルの計 26 ピクセルに対して局所的に極値であれば特徴点とする。これにより特徴点の位置とスケールを得ることができる。

入力である R 成分画像と G 成分画像に対して SIFT を適用することで特徴点を検出した画像を生成し特徴点の位置とスケールを得る。また、R 成分画像と G 成分画像の特徴点に対して、式 (2) で計算した LoG の値を、それぞれセントロメアのスコア s_{cent} 、テロメアのスコア s_{telo} とする。

また検出した特徴点の中には誤検出が含まれており、抑制しなければ染色体の検出の精度が大きく低下する。精度の向上のため LoG の計算結果から下に凸な特徴点を持った点の除去、セントロメア/テロメアとしてあり得るスケールの限定、色相からふさわしくない特徴点の除去、行って誤検出を抑制する。

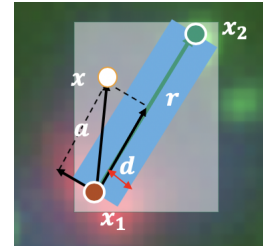


図 4 DNA の領域設定

3.3 DNA スコアの計算

3.3.1 DNA スコアの計算方法

検出された特徴点に対して、一定距離内にあるセントロメア-テロメア間、テロメア-テロメア間、セントロメア-セントロメア間を結んだ線分から一定距離内の領域（矩形領域）を作成し、その領域内の青色成分の値の平均値を DNA スコアとして取得する。

図 4 のようにセントロメア x_1 とテロメア x_2 の DNA スコアの計算を例に説明する。 x_1 から x_2 のベクトルを r 、 x_1 から探索対象 x へのベクトルを a とする。このとき DNA スコアの計算対象領域（画素の集合） S を式 (4) のように定義する。

$$S = \{x | 0 \leq l(x) \leq \|r\|, \sqrt{\|a\|^2 - (l(x))^2} \leq d\} \quad (4)$$

$$l(x) = (x - x_1) \cdot \bar{r} \quad (5)$$

次に DNA スコア s_{PAF} の定義を式 (6) で示す。ここでは計算対象領域における青成分の輝度値の平均を計算している。 $I_B(x)$ は場所 x における青成分の輝度値である。

$$s_{PAF} = \frac{1}{|S|} \sum_{x \in S} I_B(x) \quad (6)$$

3.3.2 他染色体への接続の抑制

DNA スコアを式 (6) で示したが、他染色体への接続に対してスコアを下げたり、負の値を示したりするような工夫を行うことで他染色体への接続を抑制する。

特徴点間の距離が大きい接続は他染色体に接続している可能性が高いので、これを抑制することで他染色体への接続を抑制することができる。接続する特徴点の距離が大きくなるのに応じてスコアを下げるように DNA スコアを再定義することで他染色体への接続を抑制する。式 (6) を修正した DNA スコア s_{PAF} の定義を式 (7) に示す。 d は特徴点間の距離、 k は定数である。

$$s_{PAF} = \frac{1}{|S|} \sum_{x \in S} I_B(x) \exp(-kd) \quad (7)$$

また背景部分を通る特徴点間の接続に対して負の値を与えることで他染色体への接続を抑制する。負の値を与えるために、青色画像の輝度値の平均値と標準偏差を取得し、

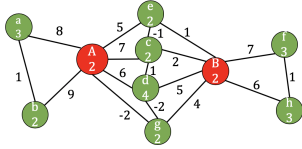


図 5 グラフの例

	A	B	a	b	c	d	e	f	g	h
A	2		8	9	7	6	5		-2	
B		2			2	5	1	7	4	6
a	8		3	1						
b	9		1	2						
c	7	2			2	1	-1			
d	6	5			1	4			-2	
e	5	1			-1		2			
f		7						3		1
g	-2	4				-2			2	
h		6						1		2

図 6 図 5 から作成した接続行列

これらを用いて青成分の輝度値を正規分布に従うように変形を行う。背景部分の輝度値に分布が偏っており DNA 部分の輝度値はピークをもたないため正規分布の軸を一定の値減少させることで背景部分を通る接続にペナルティを付与することができる。これに伴って DNA スコア s_{PAF} の定義を式 (7) を修正し式 (8) に示す。青色画像の平均値を μ_B 、標準偏差を σ_B 、全体を減少させる定数を z_{thresh} とする。

$$s_{PAF} = \frac{1}{|S|} \sum_{x \in S} \left(\frac{I_B(x) - \mu_B}{\sigma_B} - z_{\text{thresh}} \right) \exp(-kd) \quad (8)$$

3.4 グラフ分割による染色体の検出

本節では、3.2 節と 3.3 節で求めたセントロメア/テロメア/DNA の各スコアを利用してセントロメアとテロメアを接続する。

3.4.1 グラフの構築

セントロメアとテロメアの接続を行うためにグラフの構築を行う。ここではセントロメアが M 個、テロメアが N 個検出されているとする。 $M = 2$, $N = 8$ のときのグラフの例及びそこから構築された接続行列を図 5, 図 6 に示す。接続行列の要素 s_{ij} の定義を式 (9) に示す。

$$s_{ij} = \begin{cases} s_{ij}^{PAF} & (i \neq j) \\ s_i^{\text{cent}} & (i = j \text{ and } i \leq M) \\ s_i^{\text{telo}} & (i = j \text{ and } i > M) \end{cases} \quad (9)$$

次に接続の方法について考える。 i 番目のセントロメアと接続されるテロメア集合を $T_i = \{t_{ik}\}$, t_{ik} は、 i 番目のセントロメアに接続される k 番目のテロメアの ID とする。このとき i 番目のセントロメアに対する総合スコア $s_i^{\text{total}}(T_i)$ は式 (10) になる。

$$s_i^{\text{total}}(T_i) = s_i^{\text{cent}} + \sum_{j \in T_i} (s_j^{\text{telo}} + s_{ij}^{PAF}) \quad (10)$$

$$\text{s.t. } |T_i| \leq 4 \quad \forall i \quad (11)$$

$$T_i \cap T_l = \emptyset \quad \forall i, l \quad (12)$$

さらに、各セントロメアに対するテロメア集合をまとめた集合 $T = \{T_i\}$ を考えると、全体の評価関数 $s^{\text{total}}(T)$ は式 (13) となる。

$$s^{\text{total}}(T) = \sum_{i=1}^M s_i^{\text{total}}(T_i) \quad (13)$$

そして、最適なテロメア集合 T^* は式 (14) で求まる。

$$T^* = \underset{T}{\operatorname{argmax}} s^{\text{total}}(T) \quad (14)$$

最適なテロメア集合をもとにセントロメア・テロメアを接続する。式 (14) を求める近似全体最適アルゴリズムを提案する。

近似全体最適アルゴリズムは、スコア (式 (10)) の大きいセントロメアからテロメア集合を決定しながら、他染色体と接続するテロメアを組み合わせ最適に交換することで交換する双方のセントロメアの合計スコアを高くするアルゴリズムである。

まず、グラフの例を図 5 に示す。グラフから作成した接続行列 (図 6) に対して各セントロメアに対して式 (10) のスコアを産出して最大のセントロメアを探索する (図 7)。ここでは説明の簡略のために s_{PAF} のみのスコアを計算する。セントロメア A のスコアは 30, セントロメア B のスコアは 22 であるのでセントロメア A の接続対象である a, b, c, d がセントロメアの接続先になる。次にセントロメア B の接続先を考える。スコアの大きい順に接続先にしていく。テロメア f, h を接続先とし、次にテロメア d を接続先としたいがセントロメア A が接続しているため交換するか考える。図 9 のようにセントロメア B がテロメア d と接続せずにテロメア e, f, g, h と接続する場合のセントロメア A, B のスコアの和を求める、この場合 48 である。次に図 10 のようにセントロメア A がテロメア d と接続せずにテロメア d, f, g, h と接続する場合のセントロメア A, B のスコアの和を求める。この場合 51 である。なので図 10 の場合の方が全体として高いスコアとなるためテロメア d の接続先の交換をおこなう。最終的な接続の結果は図 11 のようになり、不必要なエッジを削除したグラフが図 12 になる。

4. 実験

本章では 3 章で述べた手法に基づき染色体の検出を行い、その精度を検証し考察する。

4.1 実験設定

入力として、3Gy の放射線が照射された染色体を広島大学田代聡教授らにより開発された染色法 [31] を用いて染色された画像に対して 3 章で述べた手法に基づき染色体検出

	A	B	a	b	c	d	e	f	g	h
A	2		8	9	7	6	5		-2	
B		2			2	5	1	7	4	6
a	8		3	1						
b	9		1	2						
c	7	2			2	1	-1			
d	6	5			1	4			-2	
e	5	1			-1	2				
f		7						3		1
g	-2	4				-2			2	
h		6						1		2

図 7 1 回目のスコア計算後の
接続行列

	A	B	a	b	c	d	e	f	g	h
A	2		8	9	7	6	5		-2	
B		2			2	5	1	7	4	6
a	8		3	1						
b	9		1	2						
c	7	2			2	1	-1			
d	6	5			1	4			-2	
e	5	1			-1	2				
f		7						3		1
g	-2	4				-2			2	
h		6						1		2

図 8 1 回目の選択されたセント
ロメア/テロメアを使用済み
にした接続行列

	A	B	a	b	c	d	e	f	g	h
A	2		8	9	7	6	5		-2	
B		2			2	5	1	7	4	6
a	8		3	1						
b	9		1	2						
c	7	2			2	1	-1			
d	6	5			1	4			-2	
e	5	1			-1	2				
f		7						3		1
g	-2	4				-2			2	
h		6						1		2

図 9 競合したテロメアを交換
しなかった場合の接続行列

	A	B	a	b	c	d	e	f	g	h
A	2		8	9	7	6	5		-2	
B		2			2	5	1	7	4	6
a	8		3	1						
b	9		1	2						
c	7	2			2	1	-1			
d	6	5			1	4			-2	
e	5	1			-1	2				
f		7						3		1
g	-2	4				-2			2	
h		6						1		2

図 10 競合したテロメアを交換
した場合の接続行列

	A	B	a	b	c	d	e	f	g	h
A	2		8	9	7	6	5		-2	
B		2			2	5	1	7	4	6
a	8		3	1						
b	9		1	2						
c	7	2			2	1	-1			
d	6	5			1	4			-2	
e	5	1			-1	2				
f		7						3		1
g	-2	4				-2			2	
h		6						1		2

図 11 接続行列の最終状態

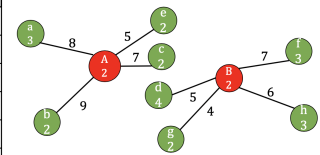


図 12 グラフの最終状態

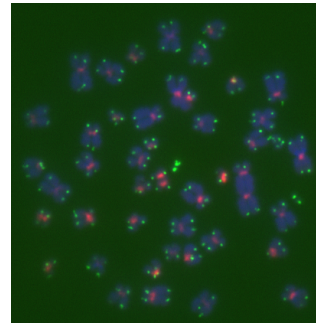


図 13 入力画像

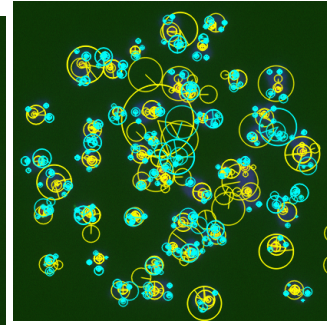


図 14 誤検出の抑制を行って
いない検出

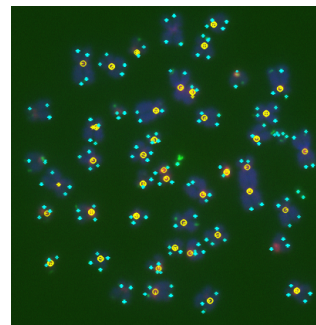


図 15 LoG, スケール, 色相に
よる抑制

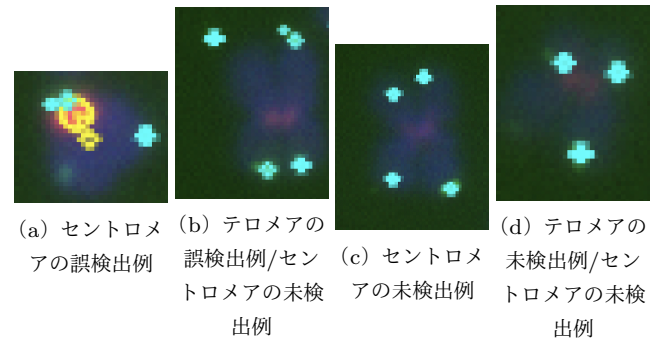


図 16 誤検出・未検出の例

を行う。本章では、テロメア/セントロメアの検出、及びグラフ分割により接続されたエッジに対して再現率と精度の評価を行う。エッジの評価の際セントロメア・テロメアの検出ミスによって起きるエッジの検出ミスは考慮しない。提案手法である近似全体最適アルゴリズムとスコアの大きい分割から順に分割していく Greedy アルゴリズムのそれぞれに再現率と精度を評価し比較を行う。

再現率 p_{recall} (式 (15)) と精度 $p_{\text{precision}}$ (式 (16)) の定義を示す。 N_{TP} を正検出数, N_{FN} を未検出数, N_{FP} を誤検出数とする。

$$p_{\text{recall}} = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FN}}} \quad (15)$$

$$p_{\text{precision}} = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FP}}} \quad (16)$$

4.2 定性的な結果

4.2.1 セントロメア/テロメア検出の定性的結果

染色体の検出の結果を示すために、まずセントロメア/テロメアの検出の結果を示す。結果は抽出した特徴点を位置とサイズと方向情報を持った点として、セントロメアの検出結果は黄色、テロメアの検出結果は水色でもとの画像に描画している。誤検出の抑制を行わず検出した結果、抑制を行って検出した結果を示す (図 13~図 15)。また、誤検出・未検出の例を図 16 に示す。

4.2.2 グラフ分割による接続の定性的結果

染色体の検出の結果を式 (10) の大きい集合から決定していく Greedy アルゴリズムと近似全体最適アルゴリズムで結果にどのような違いが出ているかを示す。一つのセントロメアからの接続はすべて同じ色で描画を行っている。各アルゴリズムの結果を比較する (図 17~図 18)。また誤

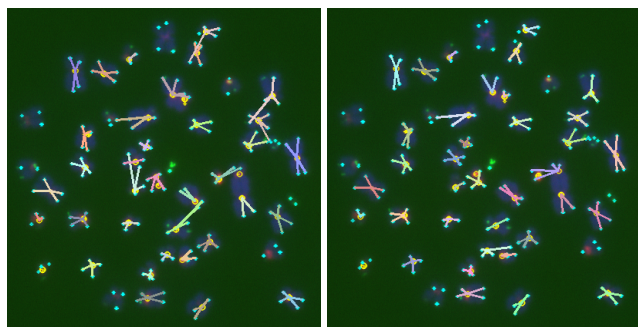


図 17 Greedy アルゴリズムを適用した結果
図 18 近似全体最適アルゴリズムを適用した結果

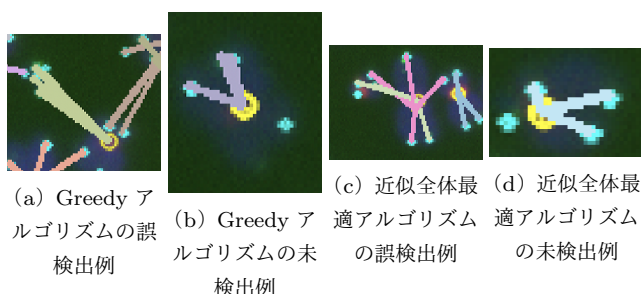


図 19 誤検出・未検出の例

表 1 セントロメア/テロメアの検出結果

	正検出	未検出	誤検出	再現率	精度
セントロメア	382	76	82	83.4%	82.3%
テロメア	1640	138	75	92.2%	95.6%

表 2 エッジの検出結果

	正検出	未検出	誤検出	再現率	精度
Greedy	1132	12	210	99.0%	84.4%
提案手法	1193	14	85	98.8%	93.3%

検出・未検出の例を図 19 に示す。

4.3 定量的な結果

4.3.1 セントロメア/テロメア検出の定量的結果

染色体画像に対してセントロメア/テロメアの検出に対する再現率 (式 (15)) と精度 (式 (16)) を表 1 に示す。

4.3.2 グラフ分割による接続の定量的結果

染色体画像に対して染色体の検出を行い、接続されたエッジに対して再現率 (式 (15)) と精度 (式 (16)) を表 2 に示す。

4.4 考察

4.3.2 節の結果から誤検出が 6.7%含まれていることがわかる。セントロメア/テロメアの検出精度の低さにより染色体の検出が不可能になることが主要な原因である。

まずセントロメアとテロメアの検出精度を比較するとセントロメアがより精度が低いことが分かる。その原因としてテロメアは円形の特徴を持った特徴点になっているがセントロメアは楕円であったり二つの円が連なった形になっ

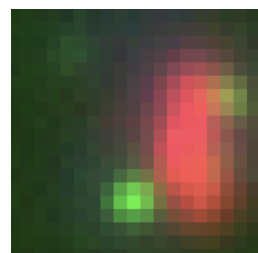


図 20 セントロメアの楕円形状

ていることが多い (図 20)。SIFT はそのような形の検出に適していないため、深層学習によりセントロメア/テロメアの特徴を学習し、高い精度で検出できると考える。

5. おわりに

本研究では、セントロメアとテロメアを検出し、検出された部位間を接続する DNA スコアを計算した。スコアを用いて、染色体画像全体に対するグラフを構築して、個別の染色体への分割をグラフ分割によって実現した。

実験ではセントロメア、テロメアの検出の再現度と検出精度の評価、及び染色体の検出の評価を行なった。現状は特徴点の誤検出や他染色体同士の部位の接続から誤検出が一定観測されている。染色体の検出の精度もいい結果とは言えないがボトムアップ手法を用いて検出を行うプログラムを実現できる可能性を示せたことは大きい。

今後は、深層学習を用いた人物姿勢推定手法である OpenPose を応用して、染色体の検出の向上及び染色体異常の分類の実装を目指す。これに向けて現状のプログラムの誤検出部、未検出部に対して修正をおこない真値を作成する機能の実装を行う予定である。

謝辞 この研究は「物質・デバイス領域共同研究拠点」の緩やかな連携の共同研究として実施されたものです。

参考文献

- [1] Y. Toyoshima, T. Tokunaga, O. Hirose, M. Kanamori, T. Teramoto, M. S. Jang, S. Kuge, T. Ishihara, R. Yoshida, and Y. Iino, "Accurate automatic detection of densely distributed cell nuclei in 3d space," *PLoS computational biology*, vol. 12, no. 6, 2016.
- [2] M. Toratani, M. Konno, A. Asai, J. Koseki, K. Kawamoto, K. Tamari, Z. Li, D. Sakai, T. Kudo, T. Satoh *et al.*, "A convolutional neural network uses microscopic images to differentiate between mouse and human cell lines and their radioresistant clones," *Cancer research*, vol. 78, no. 23, pp. 6703–6707, 2018.
- [3] Y. Sakai, S. Takemoto, K. Hori, M. Nishimura, H. Ike-matsu, T. Yano, and H. Yokota, "Automatic detection of early gastric cancer in endoscopic images using a transferring convolutional neural network," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2018, pp. 4138–4141.
- [4] 山田圭祐, 郭清蓮, 小木美恵子, 鎌田洋 *et al.*, "顕微鏡画像に基づく癌細胞計数システムの構築," 第 77 回全国大会講演論文集, vol. 2015, no. 1, pp. 657–658, 2015.

- [5] 辻幸喜, 陸慧敏, 金亨燮, 米田和恵, 田中文啓 *et al.*, “細胞領域の論理積を用いた蛍光顕微鏡画像からの血中循環がん細胞の自動検出,” *医用画像情報学会雑誌*, vol. 34, no. 4, pp. 151–155, 2017.
- [6] H. Romm, E. Ainsbury, S. Barnard, L. Barrios, J. Barquero, C. Beinke, M. Deperas, E. Gregoire, A. Koivisto, C. Lindholm *et al.*, “Automatic scoring of dicentric chromosomes as a tool in large scale radiation accidents,” *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, vol. 756, no. 1-2, pp. 174–183, 2013.
- [7] Y. Kodama, D. Pawel, N. Nakamura, D. Preston, T. Honda, M. Itoh, M. Nakano, K. Ohtaki, S. Funamoto, and A. Awa, “Stable chromosome aberrations in atomic bomb survivors: results from 25 years of investigation,” *Radiation Research*, vol. 156, no. 4, pp. 337–346, 2001.
- [8] M. S. Pearce, J. A. Salotti, M. P. Little, K. McHugh, C. Lee, K. P. Kim, N. L. Howe, C. M. Ronckers, P. Rajaraman, A. W. Craft *et al.*, “Radiation exposure from ct scans in childhood and subsequent risk of leukaemia and brain tumours: a retrospective cohort study,” *The Lancet*, vol. 380, no. 9840, pp. 499–505, 2012.
- [9] K. He, G. Gkioxari, P. Dollár, and R. Girshick, “Mask r-cnn,” in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2961–2969.
- [10] U. Iqbal and J. Gall, “Multi-person pose estimation with local joint-to-person associations,” in *European Conference on Computer Vision*. Springer, 2016, pp. 627–642.
- [11] Y. Chen, Z. Wang, Y. Peng, Z. Zhang, G. Yu, and J. Sun, “Cascaded pyramid network for multi-person pose estimation,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 7103–7112.
- [12] S.-E. Wei, V. Ramakrishna, T. Kanade, and Y. Sheikh, “Convolutional pose machines,” in *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, 2016, pp. 4724–4732.
- [13] L. Pishchulin, E. Insafutdinov, S. Tang, B. Andres, M. Andriluka, P. V. Gehler, and B. Schiele, “Deepcut: Joint subset partition and labeling for multi person pose estimation,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 4929–4937.
- [14] E. Insafutdinov, L. Pishchulin, B. Andres, M. Andriluka, and B. Schiele, “Deepcut: A deeper, stronger, and faster multi-person pose estimation model,” in *European Conference on Computer Vision*. Springer, 2016, pp. 34–50.
- [15] Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei, and Y. Sheikh, “Openpose: realtime multi-person 2d pose estimation using part affinity fields,” *arXiv preprint arXiv:1812.08008*, 2018.
- [16] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, “Realtime multi-person 2d pose estimation using part affinity fields,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 7291–7299.
- [17] E. Insafutdinov, M. Andriluka, L. Pishchulin, S. Tang, E. Levinkov, B. Andres, and B. Schiele, “Arttrack: Articulated multi-person tracking in the wild,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 6457–6465.
- [18] 下東修, “画像解析による染色体異常の自動判定,” Master’s thesis, 広島大学, 2019.
- [19] 熊谷章平 and 堀田一弘, “識別器を用いた細胞内画像からの輝点計数,” *電子情報通信学会論文誌 D*, vol. 96, no. 4, pp. 904–908, 2013.
- [20] 青木功介, 市田越子, 高林晴夫, 北来紀子, 伊川和美, and 中野慎夫, “多量な顕微鏡画像からの有核赤血球の自動検出,” *画像電子学会誌*, vol. 37, no. 5, pp. 609–616, 2008.
- [21] A. Hamano, K. Fujisaki, S. Uchida, and O. Shiku, “Stable marriage algorithm for tracking intracellular objects,” in *2013 First International Symposium on Computing and Networking*. IEEE, 2013, pp. 305–307.
- [22] M. K. Cheezum, W. F. Walker, and W. H. Guilford, “Quantitative comparison of algorithms for tracking single fluorescent particles,” *Biophysical journal*, vol. 81, no. 4, pp. 2378–2388, 2001.
- [23] M. Shimomura and K. Hotta, “Melanosome tracking using prediction by support vector regression and revision by appearance features,” in *10th International Conference on Signal Processing*, 2013, pp. 452–457.
- [24] 堀田一弘, “細胞内画像からの粒子検出と追跡,” *Medical Imaging Technology*, vol. 33, no. 3, pp. 90–96, 2015.
- [25] 橋詰明英, 鈴木隆一, 横内久猛, 堀内秀之, and 山本真司, “赤血球自動識別アルゴリズムとその評価,” *医用電子と生体工学*, vol. 28, no. 1, pp. 25–32, 1990.
- [26] 橋詰明英, 本池順, 鈴木隆一, and 山本真司, “標本内情報をを用いた適応的好中球分類アルゴリズム,” *医用電子と生体工学*, vol. 27, no. 3, pp. 149–154, 1989.
- [27] R. Girshick, “Fast r-cnn,” in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 1440–1448.
- [28] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” *arXiv preprint arXiv:1409.1556*, 2014.
- [29] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, “Semantic image segmentation with deep convolutional nets and fully connected crfs,” *arXiv preprint arXiv:1412.7062*, 2014.
- [30] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” *International journal of computer vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [31] L. Shi, K. Fujioka, J. Sun, A. Kinomura, T. Inaba, T. Ikura, M. Ohtaki, M. Yoshida, Y. Kodama, G. K. Livingston *et al.*, “A modified system for analyzing ionizing radiation-induced chromosome abnormalities,” *Radiation research*, vol. 177, no. 5, pp. 533–538, 2012.