

部分領域に基づく染色体画像解析手法

濱田智栄子[†] 阿部 亨^{††} 木下 哲男^{††}

[†] 東北大学大学院情報科学研究科 〒980-8577 仙台市青葉区片平 2-1-1

^{††} 東北大学サイバーサイエンスセンター 〒980-8577 仙台市青葉区片平 2-1-1

あらまし 顕微鏡画像を用い、染色体の種々の異常を視覚的に診断するためには、染色体領域を画像から抽出し、各領域が何番の染色体に対応するか同定を行う必要がある。これに対し筆者らは、1本の染色体領域を複数の部分領域の連なりと見なし、各部分領域を画像中で探索する処理を順に繰り返すことで、染色体領域全体の抽出・同定を高精度かつ柔軟に行う手法を提案している。本稿では、本手法を実画像に適用するための具体的な方法を説明し、さらに本手法を用い、実画像を対象に実験を行った結果を示す。

A Method for Analyzing Chromosome Image with Local Region

Chieko HAMADA[†], Toru ABE^{††}, and Tetsuo KINOSHITA^{††}

[†] Graduate School of Information Sciences, Tohoku University Sendai, 980-8577 Japan

^{††} Cyberscience Center, Tohoku University Sendai, 980-8577 Japan

Abstract To make the visual examination of a chromosome image, chromosome regions have to be extracted from the image and classified into distinct chromosome types. To improve the accuracy and flexibility in this process, we have proposed a method that regards each chromosome region as a series of subregions and iterates a search for every subregion in the image consecutively. In this article, we explain the procedures for applying the proposed method to actual chromosome images and show the experimental results of them.

1. はじめに

顕微鏡画像に基づき、染色体の種々の異常を視覚的に診断する染色体画像解析は、染色体異常に起因する疾病の予防や治療、染色体異常を誘発する物質の判定試験などにおいて重要な役割を果たす。近年、染色体の視覚的診断に関する知見が深まり、染色体画像解析の重要性が高まるにつれ、染色体画像解析に必須の処理である染色体領域の抽出・同定の自動化が以前にも増して強く求められている [1]。染色体領域を自動的に抽出・同定する手法については多くの研究がなされており、一部は製品化も行われている。しかし、十分な精度が得られる手法は未だ実現されておらず、実際の染色体画像解析に際しては、専門家による手作業で染色体領域の抽出・同定が行われる場合が多い。

筆者らは、1本の染色体領域を複数の部分領域の連なりと見なし、各部分領域を画像中で探索する処理を順に繰り返すことで、染色体領域全体の抽出・同定を高精度

かつ柔軟に行う手法を提案している。本稿では、本手法を実画像に適用するための具体的な方法を説明し、さらに本手法を用い、実画像を対象に実験を行った結果を示す。以降、2章では染色体画像解析の一般的な手順を説明し、染色体領域を抽出・同定する従来手法の問題点について述べる。3章では本稿で提案する手法の全体の処理の流れを説明する。4章では染色体の部分領域の探索方法について説明する。5章では各染色体領域の決定方法について説明する。6章では提案手法を用い、実画像を対象に実験を行った結果を示す。最後に7章でまとめと今後の課題を述べる。

2. 染色体画像解析

2.1 染色体画像解析の一般の手順

染色体は、遺伝情報を担う細胞核内の生体物質であり、ヒトの正常な細胞核1個には、常染色体44本と性染色体2本の計46本が含まれる。常染色体は22種類各2本で構成され、各種類は1, 2, ..., 22の番号で区別される。

性染色体は X, Y で表され, 正常な女性は X を 2 本, 正常な男性は X と Y を 1 本ずつ持つ. 各染色体は, 図 1 に示すように, 動原体 (cen) と呼ばれる箇所できびれており, ここで短腕 (p) と長腕 (q) に二分される [2].

顕微鏡画像に基づき, 染色体の異常を視覚的に診断する場合には, 一般に, 以下の手順で解析が行われる [3].

- (1) 染色体を染色し顕微鏡画像を撮影
- (2) 顕微鏡画像から染色体領域を抽出
- (3) 各領域が何番の染色体に対応するかを同定
- (4) 各領域の状態から染色体の異常を視覚的に診断

染色体画像の解析には, 染色体の種類 (番号) に固有の縞 (バンド) を染料で染め出し, その状態を撮影した顕微鏡画像を用いる. 染色には, いくつかの方法があるが, 染料としてギムザを使用する G 分染法 (G バンド法) が広く用いられている. 図 2 に, G 分染法で染色し撮影した顕微鏡画像の例を示す.

通常, まず, 画像から染色体領域を抽出し, 次に, 抽出された各領域の特徴に基づき, 各々が何番の染色体に対応するか同定を行う. 同じ番号の染色体でも, その大きさや濃淡は, 細胞周期での時期の違いや染色の状態等により変化するため, 染色体領域の長さ, 幅, 輝度値は画像毎に異なる (領域幅は, 異なる番号の染色体でも, 同じ画像中ではほぼ一定となる). これに対し, 染色体領域の相対的長さ, 各領域での動原体の相対的位置 (p と q の長さの比), 各領域におけるバンドの並び方 (バンドパターン) などは, 異なる画像でも染色体番号毎にほぼ一定となるため, 染色体領域の同定には, これらの特徴が用いられる [4]. 図 2 (a) 中の各染色体領域を抽出・同定し染色体番号の順に並べた結果を図 2 (b) に示す.

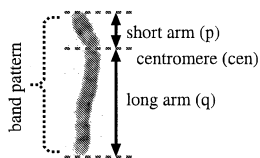


図 1 染色体の例 (5 番染色体)

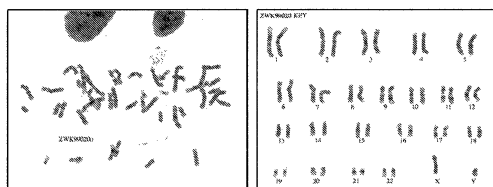


図 2 (a) G 分染法により染色した染色体の顕微鏡画像, (b) 各染色体領域を抽出・同定した結果 [5]

同定結果に基づき, 染色体の本数が通常とは異なる数的異常の診断が行われる. さらに, 同定された各領域のバンドパターンを観察すれば, 遺伝情報の欠失 (一部のバンドが欠ける), 重複 (余分なバンドが生じる), 転座 (バンドの位置が通常とは異なる) など, 染色体の構造異常の視覚的診断が可能となる.

2.2 染色体領域抽出・同定のための従来法

染色体画像解析のための処理の中でも, 個々の染色体領域の抽出・同定を自動的に行う処理については, その重要性から種々の研究がなされている. 染色体領域の同定を目的とした従来手法の多くは, 同定とは独立の前処理として染色体領域の抽出を行っており, 各染色体領域が事前に正しく抽出されていることを前提としている. しかし, 画像中の染色体領域には, 背景との区別が困難な箇所や染色体領域同士の接触・重なりが生じている箇所が存在し, 個々の染色体領域の正確な抽出は容易ではない. 染色体領域の同定手法は, 各染色体領域の (a) 全体的特徴に基づくものと (b) 部分的特徴に基づくものとに大別される [3]. 全体的特徴に基づく (a) の手法では, 各染色体ごとに, 1 本全体的のバンドパターンを参照パターンとして用意し, 抽出された染色体領域のバンドパターンと参照パターンとを比較することで同定が行われる [6], [7]. このため, 染色体領域が正確に抽出されたとしても, 領域同士が重なっている場合や染色体に構造異常が生じている場合など, 通常とは異なるバンドパターンが一部に存在する染色体領域に対しては正確な同定が困難となる. 一方, 部分的特徴に基づく (b) の手法では, 染色体領域における特徴的な一部のバンドや, p と q との長さの比率などの特徴を用い同定が行われる [8], [9]. このため, 通常とは異なるバンドパターンが一部に存在する染色体領域へも, ある程度対応することができる. しかし, (a) の手法に比べ, (b) の手法による同定の精度は低いという報告がなされている [7]. これは, 部分的特徴の安定した抽出が難しく, さらに, 同定に利用できる情報が (a) の手法に比べ少ないためと考えられる.

3. 部分領域に基づく染色体領域の抽出・同定

染色体領域の同定を目的とした従来手法には以下の問題があった.

- 個々の染色体領域が正確に抽出されていないと同定が困難
- 染色体の構造異常や領域同士の重なり等により, 通常とは異なるバンドパターンが領域の一部に生じている場合は正確な同定が困難

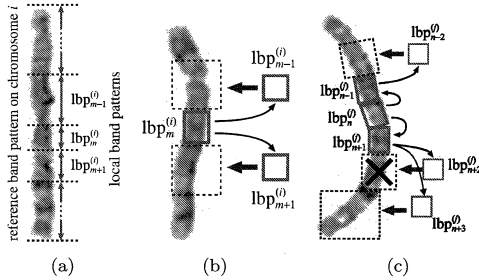


図3 部分領域を単位とした染色体領域の抽出・同定

これら従来手法における問題の解決を図り、染色体領域の高精度かつ柔軟な抽出・同定を実現するために、筆者らは、染色体領域の抽出・同定を部分領域単位で行う手法を提案している[10]。この手法では、図3(a)に示すように、抽出・同定を行う際のテンプレートとして、染色体の番号毎に事前に用意した参照バンドパターンを各々複数の部分に分割したものをを用いる。以下では、これを部分バンドパターンと呼び、 i 番染色体の m 番部分バンドパターンを $lbp_m^{(i)}$ で表す。

まず、部分バンドパターンをいくつかを選択し、図3(b)に示すように、これらをテンプレートとして用いて、類似する箇所(部分領域)を画像中で探索する。部分バンドパターン $lbp_m^{(i)}$ に類似する部分領域を検出した場合、次に、その近傍に探索範囲を設定し、隣接する部分バンドパターン $lbp_{m-1}^{(i)}, lbp_{m+1}^{(i)}$ に類似する部分領域を探索する。この反復により、最初に検出した箇所を起点に、隣接部分バンドパターンに類似する部分領域を順に探索し、1本の染色体領域全体を芋蔓式に抽出・同定する。さらに、異なる箇所を起点に探索を繰り返すことで、全染色体領域の抽出・同定を行う。

一連の過程で、ある部分バンドパターンに類似する部分領域の探索に失敗した場合、その箇所では、領域の重なりや染色体の構造異常により、通常とは異なるバンドパターンが生じている可能性が高いと考えられる。そこで、本手法では、状況に応じて次の探索範囲(画像中の領域)や探索対象(部分バンドパターン)を変更し、このような箇所への対応を図る。例えば、図3(c)に示すように、 $lbp_{n+1}^{(j)}$ の近傍で $lbp_{n+2}^{(j)}$ の探索に失敗した場合、探索範囲を変更したり、探索対象を $lbp_{n+3}^{(j)}$ に変更する等の制御を行う。

以上のアプローチにより、本手法では以下の効果が期待できる。

- 部分領域単位の探索により抽出・同定を同時に行い、さらに、探索結果を相互に利用することで、様々な状態にある染色体領域の高精度・効率的な抽出を実現。

- 探索範囲・対象を状況に応じて制御し、通常とは異なるバンドパターンを示す箇所を除きながら領域の特徴を統合することで、染色体領域の高精度・柔軟な同定を実現。

4. 染色体の部分領域の探索

部分領域を探索する際には、まず各 $lbp_m^{(i)}$ をテンプレートとして用い、対象画像中でテンプレートを移動・回転させながら、各箇所を対象画像とテンプレートとの輝度値の相違度を計算する。各箇所 (x, y) でテンプレートを回転させ、相違度が最も低くなる回転角度を (x, y) での回転角度、またそのときの相違度を (x, y) の相違度とする。対象画像中で相違度が最も低い箇所をテンプレートに対応する部分領域の候補とする。

実画像における染色体領域の大きさ(特に長さ)や輝度値は画像毎に異なる。また輝度値に関しては、1本の染色体領域内で部分的に変動している場合もある。このため、探索用のテンプレートと対象画像中の染色体領域を正しく比較できるように、両者の長さや輝度値を揃える処理が必要となる。

4.1 テンプレートの長さ・幅の決定

4.1.1 染色体領域長の正規化

染色体領域長の正規化に際しては、輝度値に基づき2値化した画像中で染色体領域幅 W と染色体領域長の総和 L を求め、この値を正規化のための基準に用いる。

1枚の画像中では、異なる番号の染色体でも、領域幅がほぼ同じ値 W となる。したがって、各染色体領域を幅が一樣な帯状の領域とみなせば、画像中での染色体領域の総面積 S は $L \times W$ で近似できる。そこで提案手法では、2値化した画像中、領域に対応する画素数の総数を S に、各箇所求めた領域幅の最頻値を W に用い、 $L = S/W$ により L を決定する。

テンプレートの作成に用いた画像 I_R と対象画像 I_S で領域幅 W_R, W_S と領域長の総和 L_R, L_S を求め、テンプレートの幅を W_S に、テンプレートの長さを L_S/L_R 倍し、テンプレートと対象画像中の染色体領域の大きさを揃える。

4.1.2 染色体領域幅の推定

以下では、2値化された対象画像中、領域に対応する画素を p_r 、背景に対応する画素を p_b で表す。染色体領域の幅は、1枚の画像中ではほぼ一定となるため、本手法では、領域境界上の各 p_r で領域幅を求め、得られた領域幅の最頻値を W に用い、対象画像中での染色体領域の幅 W を推定する。

まず、図4(a)に示すように、領域境界(p_b と隣接する箇所)の各 p_r で、 $\pi/4$ 毎8方向($t=0, 1, \dots, 7$)に、

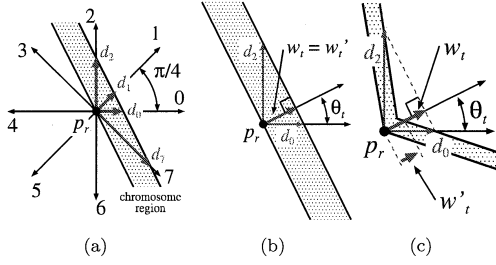


図4 境界 p_r での領域幅

他方の領域境界までのユークリッド距離 d_t を求める。二つの距離 d_t, d_{t+2} を用いると、領域幅 w_t とその方向 θ_t は、式 (1), (2) で得ることができる。

$$w_t = d_t \times d_{t+2} / \sqrt{d_t^2 + d_{t+2}^2} \quad (1)$$

$$\theta_t = \tan^{-1}(d_t/d_{t+2}) + t \times \pi/4 \quad (2)$$

ここで、 $d_8 = d_0, d_9 = d_1$ とし、計算の安定性のため d_t, d_{t+2} 共に閾値 Td 以上の場合のみ w_t, θ_t を求める。次に、 θ_t 方向に、 p_r から他方の領域境界までの距離 w'_t を求める。図 4 (b) に示すように、領域境界が直線で互いに平行な箇所では、 w_t, w'_t 共に正しい領域幅に一致する。しかし、図 4 (c) に示すように、領域境界が屈曲している箇所や平行でない箇所では、両者は一致せず、 w_t, w'_t が正しい領域幅になるとは限らない。そこで本手法では、両者の差 $e_t = |w_t - w'_t|$ が閾値 Te 以下となる w'_t のみを信頼できる領域幅と見做す。一つの p_r に対し、信頼できる w'_t が複数得られた場合（複数の t で $e_t \leq Te$ となる場合）、 e_t が最小の w'_t を p_r での領域幅として選択する。

領域境界に位置する各 p_r で選択した w'_t の度数分布を求め、 w'_t の最頻値を対象画像中での染色体領域幅 W とする。このように、提案手法では、信頼できる w'_t のみに基づいた度数分布を用いることで、 W の安定した推定を図っている。

4.2 対象画像での輝度値の局所的な補正

部分領域を探索する際に使用する相違度には、対象画像とテンプレートとの輝度値の差の2乗平均 (MSE: mean squared error) を用いることが考えられる。図 5 (a) に示すように、 (x, y) でテンプレートを θ 回転させたときの MSE $e^2(x, y)$ は式 (3) で求められる。

$$e^2(x, y) = \frac{1}{UV} \sum_{u=0}^{U-1} \sum_{v=0}^{V-1} (Is(x', y') - T(u, v))^2 \quad (3)$$

$$x' = x + u \cos \theta - v \sin \theta \quad (4)$$

$$y' = y + u \sin \theta + v \cos \theta \quad (5)$$

ここで、 Is, T は、対象画像とテンプレートの輝度値を

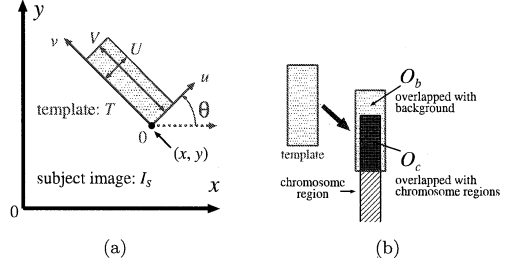


図5 テンプレートによる部分領域探索

各々表す。 U, V はテンプレートの幅と長さを表す。

染色体領域の輝度値は、染色の状態等により画像毎に異なるため、提案手法では、より安定した探索の実現を図るために、対象画像中の輝度値を局所的に線形変換し、補正された輝度値に基づき計算された MSE $\tilde{e}^2(x, y)$ を相違度に用いる。補正された輝度値に基づき計算された $\tilde{e}^2(x, y)$ は次式で与えられる。

$$\tilde{e}^2(x, y) = \frac{1}{UV} \left(\tilde{E}_b^2(x, y) + \tilde{E}_c^2(x, y) \right)^2 \quad (6)$$

$$\tilde{E}_b^2(x, y) = \sum_{(u, v) \in O_b} (Is(x', y') - T(u, v))^2 \quad (7)$$

$$\tilde{E}_c^2(x, y) = \sum_{(u, v) \in O_c} (\alpha Is(x', y') + \beta - T(u, v))^2 \quad (8)$$

ここで、 $\tilde{E}_b^2(x, y), \tilde{E}_c^2(x, y)$ は領域 O_b, O_c で各々計算された対象画像とテンプレートとの輝度値の差の2乗和を表す。図 5 (b) に示すように、テンプレートは対象画像中の背景に重なる箇所 O_b と染色体領域に重なる箇所 O_c の2つの領域に分けられる（背景か染色体領域かは対象画像を二値化したものから判定する）。対象画像中の背景と重なる箇所 O_b では、補正を行わない Is をそのまま用いて輝度値の差の2乗和 $\tilde{E}_b^2(x, y)$ を求め、染色体領域と重なる箇所 O_c では、 Is を線形変換した値 $\alpha Is(x', y') + \beta$ を用いて $\tilde{E}_c^2(x, y)$ を計算する。 α, β は、 Is を補正するための値であり、各 (x, y) で $\tilde{E}_c^2(x, y)$ が最小となるよう、式 (9), (10) から決定する（ただし、 $\alpha \leq 0$ となる箇所は除外する）。

$$\alpha = \frac{|O_c| \sum_{(u, v) \in O_c} Is(x', y') T(u, v) - \sum_{(u, v) \in O_c} Is(x', y') \sum_{(u, v) \in O_c} T(u, v)}{|O_c| \sum_{(u, v) \in O_c} Is^2(x', y') - \left(\sum_{(u, v) \in O_c} Is(x', y') \right)^2} \quad (9)$$

$$\beta = \frac{1}{|O_c|} \left(\sum_{(u, v) \in O_c} T(u, v) - \alpha \sum_{(u, v) \in O_c} Is(x', y') \right) \quad (10)$$

ここで、 $|O_c|$ は領域 O_c 内の画素数を表す。

このように、提案手法は 1 本の染色体領域内で部分的に輝度値が変動している場合へも対応可能となる。さらに、提案手法では O_b と O_c を区別し、 O_c でのみ対象画像の輝度値を補正することで、部分領域探索精度の向上を図っている。

5. 各染色体領域の決定

各染色体領域を探索する際には、まずはじめに lbp をいくつか選択し、各 lbp に類似する部分領域を画像全体で探索する。部分バンドパターン $lbp_n^{(i)}$ に類似する部分領域を検出した場合、次にその近傍に探索範囲を設定し、隣接する $lbp_{n+1}^{(i)}$ に類似する部分領域を探索する。この反復により、最初に検出した部分領域を起点に、隣接する部分バンドパターンに類似する部分領域を順に探索し、1 本の染色体領域決定する。

隣接する lbp を探索する際の手順を以下に述べる。

5.1 探索範囲の設定

$lbp_n^{(i)}$ に対し検出された部分領域から、 $lbp_{n+1}^{(i)}$ に対応する部分領域を探索する際の探索範囲を設定する。検出された個々の部分領域の状態に応じた探索を実現するために、以下のように探索範囲 $SA_{n+1}^{(i)}$ を設定する。なお、以下では図 6 (a) に示すように、 $lbp_n^{(i)}$ の 4 隅を $UL_n^{(i)}$, $UR_n^{(i)}$, $DL_n^{(i)}$, $DR_n^{(i)}$ で表す。

- 図 6 (b) に示すように、 $lbp_n^{(i)}$ に対応する部分領域を検出した際の $UL_n^{(i)}$ と $UR_n^{(i)}$ の中点を求める。
- 得られた中点を中心として、半径 $K \times V_n^{(i)}$ 、中心角 $\left(\theta_n^{(i)} + \frac{\pi}{2}\right) \pm \delta\theta$ の扇形の領域を求める。
- 求めた扇形に外接する矩形を決定し、これを探索範囲 $SA_{n+1}^{(i)}$ とする。

5.2 $lbp_{n+1}^{(i)}$ に対応する部分領域の探索

図 7 に示すように、探索範囲 $SA_{n+1}^{(i)}$ 内で $lbp_{n+1}^{(i)}$ に対応する部分領域を探索する。MSE を計算する際には、4 章で述べた領域長の正規化と輝度値の補正を行う。ただ

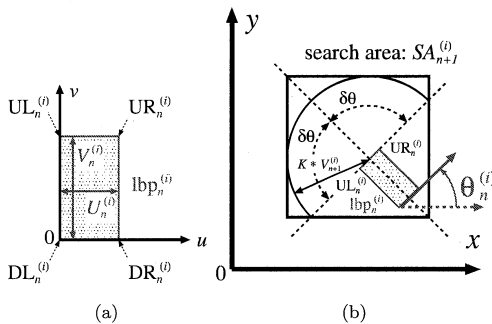


図 6 (a) $lbp_n^{(i)}$ の 4 隅, (b) 探索範囲の設定

し、探索は染色体領域のみで行い (染色体領域は画像を 2 値化し決定)、MSE を求めるためにテンプレートを回転する際には回転範囲を $\left(\theta_n^{(i)} + \frac{\pi}{2}\right) \pm \delta\theta$ とする。探索の際に各箇所 で計算された MSE が小さい順に N 箇所を $lbp_{n+1}^{(i)}$ に対応する部分領域の候補とする。

さらに、 $lbp_{n+1}^{(i)}$ に対応する部分領域と $lbp_n^{(i)}$ に対応した部分領域との位置関係を考慮することにする。 $lbp_n^{(i)}$ に対応した部分領域と $lbp_{n+1}^{(i)}$ に対応した部分領域の候補間の距離 l を $UR_n^{(i)}$ と $DR_{n+1}^{(i)}$ 間の距離とし、この距離 l が閾値 γ 以上の場合には候補から除く (図 8)。

5.3 $lbp_{n+1}^{(i)}$ に対応する部分領域の決定

$lbp_{n+1}^{(i)}$ に対応する部分領域を決定する際には、 $lbp_n^{(i)}$ に類似する部分領域と $lbp_{n+1}^{(i)}$ に類似する部分領域間で、より正確に領域を決定できるよう微調整を行う。微調整では、MSE が $lbp_n^{(i)}$ と $lbp_{n+1}^{(i)}$ の 2 つをあわせた領域 $R_{n,n+1}$ でより小さくなるよう、 $lbp_n^{(i)}$ と $lbp_{n+1}^{(i)}$ のパラメータ (参照バンドパターン中で lbp がそれぞれ担当する範囲と回転角度) を各々調整する (図 9)。

この微調整で $R_{n,n+1}$ での MSE が最も小さくなる候補を $lbp_{n+1}^{(i)}$ に対応する部分領域とする。

6. 実験

提案手法を用い、実画像を対象に各染色体領域を決定する実験を行った。

6.1 実験に使用した染色体画像

実験には、ZooWeb と The Wisconsin State Laboratory of Hygiene により公開されている染色体画像を利用

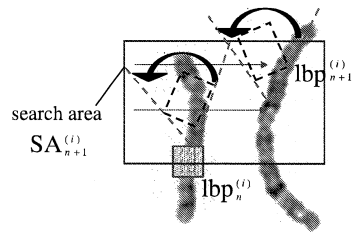


図 7 探索範囲 $SA_{n+1}^{(i)}$ 内で $lbp_{n+1}^{(i)}$ に対応する部分領域を探索

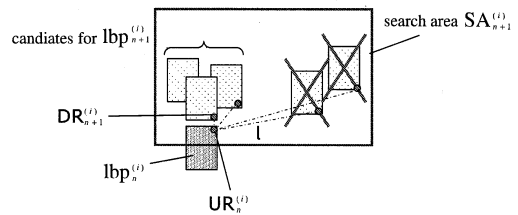


図 8 $lbp_n^{(i)}$ の検出箇所との位置関係を考慮

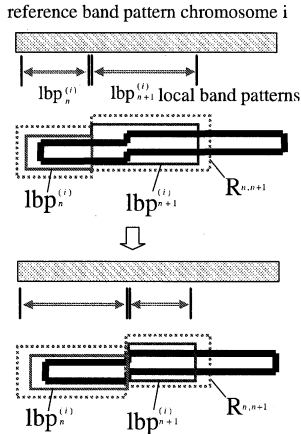


図9 パラメータの微調整

した [5]. この画像の例を図 2 に示す ((a) は原画像であり, (b) はその同定結果である). 同定結果の画像 (ImgK で表す) では, 個々の染色体領域が一部手動で分離・統合され, 染色体番号順に並べられている. ImgK は, 女性 (Y 染色体を含まないもの) 19 枚, 男性 (Y 染色体を含むもの) 12 枚の計 31 枚である. 31 枚中, 正常なものが 9 枚, 数的異常を生じているものが 22 枚 (染色体数 45 本のもの 2 枚, 47 本のもの 20 枚) あり, 構造異常を生じているものは無い. 全 31 枚の ImgK から, 探索用のテンプレート (部分バンドパターン) を作成するための画像を 16 枚選択し, 残りの 15 枚の ImgK を探索対象画像として用いた.

6.2 実験に用いたテンプレート

探索用のテンプレートは, 31 枚の ImgK の半数 16 枚から作成した. まず, ImgK から抽出した各領域の芯線を決定し (図 10 (a), (b)), 次に, 芯線に直交する方向で輝度値の平均を求め (図 10 (c)), 得られた平均値を直線的に配置することで屈曲を補正し輝度値を抽出した (図 10 (d)). これを全 16 枚・各染色体について行い, 全 16 枚各番号ごとに輝度値を平均し参照バンドパターン ImgT を作成した. さらに, ImgT を図 11 に示すように分割し, 探索に使用する lbp を作成した. 探索の起点となる lbp は図 11 中*で示した lbp とする.

6.3 実験結果

実験では 1, 2, ..., 5 番染色体領域の決定を行った (従って, 実験では各探索対象画像中で 1, 2, ..., 5 番染色体 \times 2 本の計 10 本の染色体領域を決定する). なお, 起点となる lbp に対応する箇所は, 予め正しい位置を指定しておき, 用意した lbp をすべて使用し, 染色体領域を正しく決定できた場合を染色体領域の決定が成功した

こととする.

領域長の正規化に用いる L (画像中の領域長の総和) は, 文字・記号を削除した ImgK (図 12 (a)) を 2 値化し, 孤立点を除去した画像 (図 12 (b)) を対象に, 領域の幅 W の推定 (閾値は $Td = 3$ 画素, $Te = 2$ 画素とした) と領域の総画素数 S の計測を行った結果から求めた. 探索範囲を設定する際は, $\delta\theta = 90^\circ$, $K = 2$ とした (つまり扇形の半径は $lbp_{n+1}^{(i)}$ の長さの 2 倍とする). また, 正規化で用いる 2 値化された画像 (図 12 (b)) は, 探索の際に染色体領域かどうかを判定する際にも使用する.

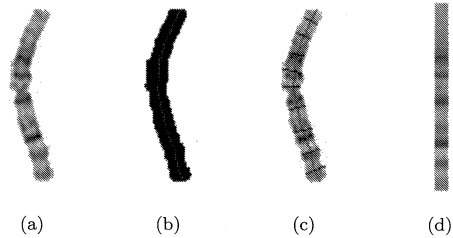


図 10 (a) 濃淡画像, (b) 抽出された領域・決定された芯線, (c) 輝度値の決定, (d) 生成された染色体領域

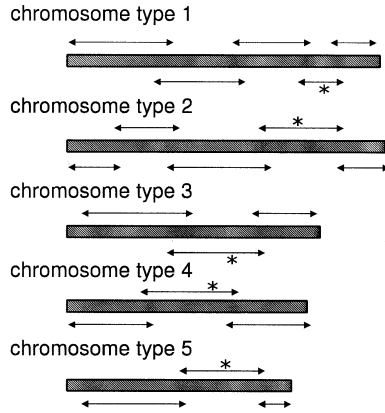


図 11 部分領域の探索に用いたテンプレートの例

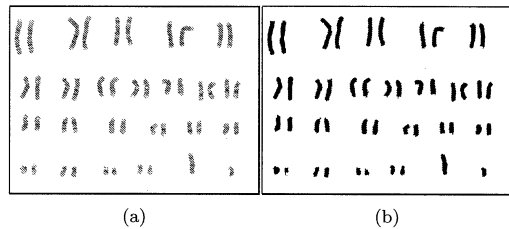


図 12 (a) 文字・記号を削除した ImgK, (b) 2 値化後に孤立点を除去した結果

図 13, 図 14 に決定された 1 番染色体領域の例を示す。また, 同じ lbp を用い, 原画像に対し同様に 1 番染色体領域を決定した結果を図 15 に示す。図 16 は決定された 2~5 番染色体領域の例である。

今回探索した染色体の本数は, 各画像 10 本 × 15 枚でのべ 150 本, うち染色体領域の決定に成功した染色体の本数はのべ 101 本であった。従って, 染色体領域の決定の成功率を, 染色体領域の決定に成功した染色体の本数/探索した染色体の全本数とすると, 成功率は約 67% である。

次に決定に失敗した例を図 17 に示す。決定に失敗した原因を調査したところ, 染色体領域が極端に屈曲している場合は誤った箇所を検出していた。また, 1 カ所でも lbp の検出に失敗すると, その先の lbp の検出も失敗していた。これは, 1 つ前の lbp の検出結果の情報をもとに, 次の lbp に対応する部分領域の探索を行う際の探索範囲を決定しているためである。今後, 屈曲等で lbp に対応する部分領域の検出が困難な箇所への対応や, lbp に対応する部分領域の検出に失敗した場合の探索範囲の設定法についてもさらに検討する必要がある。

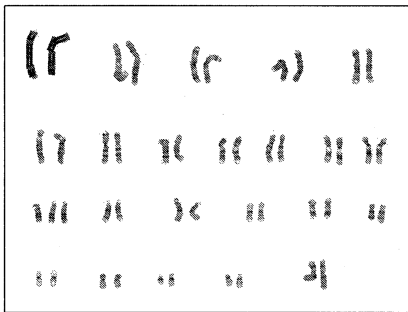


図 13 1 番染色体領域の決定例

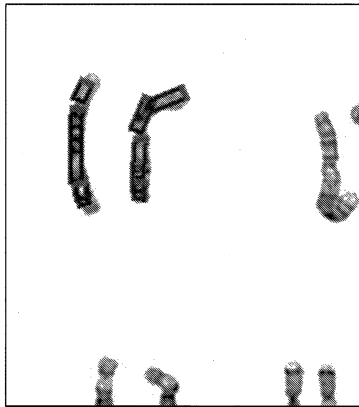


図 14 1 番染色体領域の決定例 (拡大)



図 15 1 番染色体領域の決定例 (原画像)

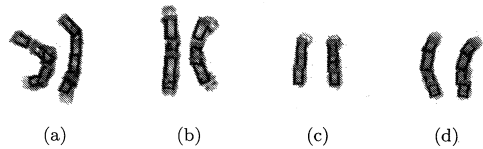


図 16 領域の決定例 (a) 2 番染色体, (b) 3 番染色体, (c) 4 番染色体, (d) 5 番染色体



図 17 領域の決定に失敗した例

7. おわりに

本稿では, 1 本の染色体領域を複数の部分領域の連なりと見なし, 部分領域単位での探索を行うことで染色体領域全体を効果的に抽出・同定する手法を実画像に適用するための具体的な手法について説明した。さらに, 本手法を用い, 実画像を対象に実験を行った結果を示した。

今後はより精度の高い探索の実現を目指し, 以下の点を中心にさらに検討を進める予定である。

- ある lbp に対応する部分領域の検出に失敗した場合, 次の探索範囲や探索対象 (lbp) を変更するといった制御が必要となる。これを実現するためには, 部分領域の検出に失敗したかどうかを判定する方法について検討する必要がある。

- 探索範囲を変更する場合、新たな探索範囲をどのように設定するかについて検討する必要がある。

- 最終的に決定された1本の染色体領域が正しいかどうか判定する方法について検討する必要がある。

謝辞 本研究を進めるにあたり、御討論頂きました(株)日本遺伝子研究所 島田 裕充 氏、日本アイ・ビー・エム (株) 大庭 信之 博士に感謝致します。

文 献

- [1] A. Carothers and J. Piper: "Computer-aided classification of human chromosomes: a review", *Statistics Computing*, **4**, 3, pp. 161-171 (1994).
- [2] 新川, 阿部: "遺伝医学への招待", 南江堂, 改訂第3版 (2005).
- [3] J. Graham and J. Piper: "Automatic karyotype analysis", *Meth. Mol. Biol.*, **29**, pp. 141-185 (1994).
- [4] D. G. Harnden and H. P. Klinger Eds.: "ISCN1985: An international system for human cytogenetic nomenclature (1985)", S. Karger AG (1985).
- [5] "ZooWeb - Karyotypes home page", <http://worms.zoology.wisc.edu/zooweb/Phelps/karyotype.html>.
- [6] J. Piper and E. Granum: "On fully automatic feature measurement for banded chromosome classification", *Cytometry*, **10**, 3, pp. 242-255 (1989).
- [7] Q. Wu, Z. Liu, T. Chen, Z. Xiong and K. R. Castleman: "Subspace-based prototyping and classification of chromosome images", *IEEE Trans. Pattern Anal. Machine Intell.*, **14**, 9, pp. 1277-1287 (2005).
- [8] F. C. A. Groen, T. K. Kate, A. W. M. Smeulders and I. T. Young: "Human chromosome classification based on local band descriptors", *Pattern Recognit. Lett.*, **9**, 3, pp. 211-222 (1989).
- [9] M. Moradi and S. K. Setarehdan: "New features for automatic classification of human chromosomes: A feasibility study", *Pattern Recognit. Lett.*, **27**, 1, pp. 19-28 (2006).
- [10] 山口, 濱田, 島田, 浅野, 酒井, 大庭, 佐藤, 阿部, 木下: "部分バンドパターンに基づく染色体領域の抽出・同定", 第10回画像の認識・理解シンポジウム MIRU2007 講演論文集, pp. 1384-1389 (2007).