

Determination of Overlapped Region of Chromosome Images for Disentanglement and Karyotyping.

A.Vishnu Priyesh

PG Student (M E Communication Systems)
Department of Electronics and Communication
Engineering
Sri Venkateswara College of Engineering.

Mrs.D.Menaka

Associate Professor
Department of Electronics and Communication
Engineering
Sri Venkateswara College of Engineering.

Abstract— the objective of the proposed work is to separate the touching and overlapped chromosome images in chromosome images. Chromosome anomalies are of two types Numerical and structural. The process of determining these abnormalities, by a cytogeneticist, is realized by investigating a number of chromosome images. This is a time consuming procedure and often may lead to an error. Karyotype is the main method for diagnosing major genetic disorders. In karyotyped image, the chromosomes are arranged according to the International System for Cytogenetic Nomenclature (ISCN) classification. The initial step is to separate the overlapping and touching chromosome images. Segmentation is an important step for identifying the region of interest. Initially the input grey scale image of chromosome was subjected to various pre-processing techniques and then watershed algorithm technique was applied over the processed image to obtain the segmented chromosomes. Active Contour Technique was adopted to segment the chromosome as watershed segmentation had poor efficiency, alteration in algorithm was carried out to effectively segment the chromosomes as they are non-rigid structures and again the overlapped region is detected with the help of active contour.

Keywords— Chromosomes; Active Contour; Cytogenetic; ; Segmentation.

I. INTRODUCTION

Birth defect is an universal problem. Comprising of abnormalities in new born baby's structure, function or body metabolism which leads to physical and mental dysfunctions and can be fatal sometimes. One of the main causes of this is chromosomal disorders. Chromosomes is made up of supercoiled DNA and its associated proteins. A chromosome is identified with 24 classes where 22 pairs of chromosomes are autosomes. Each autosome is assigned a number from 1-22 and the 23rd pair is the sex

Chromosome referred as XX or XY. Karyotyping analysis is an important screening and diagnostic procedure for detecting several genetic diseases or chromosomal anomalies. A chromosome anomaly depicts a typical number of chromosomes or a structural abnormality in one or more chromosomes. The numerical anomaly is a variation in number of chromosome. Structural anomaly is the breakage and loss of a portion of chromatid arm or a reunion of the arm at different location on the same chromosome or on a different chromosome. Chromosome anomalies usually

occur when there is an error in cell division following meiosis or mitosis operation.

Occasionally, chromosomal material is lost or rearranged during cell division or during the formation of gametes leading to severe abnormalities. Thus, it originates the need to analyze the number and shape of chromosomes to observe the possibility of origin of various diseases. Thus, it helps in calculation of the total number of chromosomes present in a karyotype and also separate the touching or overlapping chromosomes. Such images have to be further processed to separate them to form single chromosome images, with which only karyotype can be formed.

II. KARYOTYPING

The preparation and study of karyotype is a part of cytogenetics. In it, a fresh sample of human blood is taken and then planted for cell division. It is then stained with materials such as Giemsa. Besides having a property of passing hereditary information chromosomes have many other properties. One of them is it being highly colorable. The coloring property of chromosomes is of great help to us in this process as the stage at which the chromosomes are best colored and maximum clarity is present we stop the cell division. Usually, this is the metaphase stage of cell division. After acquiring the images the chromosomes are arranged in a pre-defined pattern based upon their shape, length and the location of centromere to form the karyotype. This is the standard pattern for the study of chromosomes.

A. Purpose of karyotyping

The study of karyotype is the basis of detection of various diseases. The basic genetic structure of life and the traits to be passed on to offspring can be known by karyotyping. The given karyotype is compared with an ideogram (standard karyotype) to know the possibility of a disease or the sex of the human being. Further processing is needed for this purpose such as labelling of chromosomes to their pair number and then analyzing their number, shape and length during comparison. The genetic abnormalities in the chromosomes also have to be considered in this process.

III. LITERATURE SURVEY

A. Overlapped region detection using curvature Function

In earlier known works, the overlapped region is detected using curvature function. Curvature function is nothing but rate of change of slope with respect to length. Concavity and convexity are the two concepts involved in curvature function, concave points are obtained with the help of curvature function the angular difference between the successive nodes gives the curvature function

The initial scan process starts at 't' position which is slightly deviated from zero. This deviation gives the length t+s and helps the curvature to be constant through all non-zero points. The scanning of closed n link chain with pixel points and the coordinates are traced

The concave property of a region is used to find the Possible real concave points along a contour. The concavity of a point depends mainly on its neighboring points. The property is explained by the current and previous values. The separation lines are obtained by considering all possible pairs of high concave points. The initial set of possible separation lines is obtained by taking the possible combinations of points, excluding the reflexive pairs. The separation lines are traced by pixels with non-zero values and zero values as background. The initial point of the image should be considered by specifying the row and column coordinates and the initial search direction should be specified for the next connected pixel. The other parameters to be considered are the connectivity and the direction to trace the connected pixel for obtaining the separation lines. The hypothesis is determined from these separation lines. The hypothesis is based on geometric information which is obtained as the interesting points on the image. This is further used to separate the overlapping and touching portion in the image.

B. Geometric Approach

In this work, a geometric-based method is used for automatic detection of touching and overlapping chromosomes and separating them. The proposed scheme performs segmentation in two phases. In the first phase, chromosome clusters are detected using three geometric criteria, and in the second phase, chromosome clusters are separated using a cut-line.

C. MOGAC

In this approach Active contour technique is combined with the geometric approach for detecting overlapped regions of chromosomes

IV. PROPOSED APPROACH

A. Introduction to split bregman active contour method

Segmentation and object extraction is one of most important tasks in image processing and computer vision. Many of the most general and effective segmentation methods can be written as variational/PDE based models. This category of varying models has been shown to be very effective in many applications, especially in the processing

and analysis of medical images. While there are many disparate approaches to image segmentation, here it is done by using recently proposed methods which can be cast in the form of totally convex optimization problems. This convexity property allows segmentations to be computed using fast elliptic solvers.

Several different varying frameworks for image segmentation have been proposed, most of which fall into one of two categories. The first such category we will discuss is the geodesic active contour (GAC)/snakes model. Originally proposed snakes based segmentation identifies objects using an edge detector function, which takes on small values near boundaries (e.g. where the image gradient is large) and large values (typically near unity) where the image is smooth. Following the GAC method of Caselles, Kimmel, and Sapiro, this is accomplished using a curve evolution procedure by solving

$$\min_c \int_c^\infty g(\Delta f) ds \quad (1)$$

Where C represents the closed boundary curve, ∇f is the image gradient. The function g is the non-negative edge detector function. Models of this type have been studied in the context of segmentation and feature extraction by Kichenassamy et al. One common choice for the edge detector is

$$g(\epsilon) = \frac{1}{1+\beta|\epsilon|^2} \quad (2)$$

Where β is a parameter that determines the detail level of the segmentation. It has been found that, by identifying a curve lying along edges in an image, model (1) extracts relevant semantic features from an image.

The calculus of variations then allows us to find the Euler-Lagrange equation for (1) and the minimization is carried out using the following curvature-minimizing gradient flow:

$$C_t = (g_k - \langle \nabla g, N \rangle) N \quad (3)$$

Where κ represents the curvature and N is the unit normal to the curve C. various numerical schemes have been proposed for computing this gradient flow. One of the most simple and versatile schemes is the level set method of Osher and Sethian In this method, the curve C is represented using an implicit level set function, ϕ . The curve evolution is then solved using the flow

$$\phi_t = \left(g \nabla \cdot \frac{\nabla \phi}{|\nabla \phi|} - \langle \nabla g, \nabla \phi \rangle \right) |\nabla \phi| \quad (4)$$

Models of the form (1) have also been used for the related problem of reconstructing a surface from unorganized data, where g is replaced by the distance to the unorganized data set. Given a set of points $\{x_i\}$ lying on some smooth surface, we wish to reconstruct an approximation of that surface. This problem is very difficult because it is, in general, ill-posed. Also, any solution to this problem must be able to

handle a wide range of topological and geometric surface configurations. Explicit representation methods such as NURBS rely on a parameterization of the surface, which requires a significant amount of a priori knowledge of the surface topology. Methods relying on triangularizations and Voronoi diagrams are limited to reconstructing piecewise linear surface, and become impractical if dynamic surface evolution is required. A very nice approach that does not suffer from these difficulties is to use an implicit level-set representation. This can be accomplished by defining g to be the distance function of the set $\{x_i\}$, and then minimizing the corresponding energy (1). As a result, this modified GAC model computes a surface that lies along the minimal contours of this distance function (i.e. a surface that passes close to the points $\{x_i\}$).

B. The Split Bregman Method: A General L1 Minimization Technique

The Split Bregman method is a technique for solving general L1-regularized problems of the form

$$arg \min_{\mu} |\Phi\mu| + \frac{\mu}{2} \|A\mu - f\|^2 \tag{5}$$

Where Φ and A are (possibly singular) linear operators. For example, choosing $\Phi = \nabla$ and $A = I$ yields the ROF model. There is a large literature on techniques for solving (5). Many techniques approach the problem by either solving a regularized form of (5) directly, or by attacking the differentiable “dual formulation” of the problem, which requires the enforcement of linear inequality constraints. In, the Split Bregman method was introduced for solving (5). This method has the advantage that it does not require regularization, continuation, or the enforcement of inequality constraints. Furthermore, the technique has been shown to be an extremely efficient solver for L1 regularized denoising problems, as well as a large class of problems from compressed sensing. The Split Bregman method works by “de-coupling” the L1 and L2 terms in (5), using a splitting originally introduced in Rather than solve (5) directly, we introduce the auxiliary variable $d \leftarrow \Phi u$. The problem (5) then becomes

$$arg \min_{\mu} |\Phi\mu| + \frac{\mu}{2} \|A\mu - f\|^2 \tag{6}$$

To solve this constrained problem, we convert it to an unconstrained problem by introducing a quadratic penalty function.

$$arg \min_{\mu} |d| + \mu \|A\mu - f\|^2 + \frac{\lambda}{2} \|d - \Phi u\|^2 \tag{7}$$

This formulation of the problem is very advantageous because the unconstrained problem (7) can be solved using a simple alternating minimization scheme. The first step of this alternating scheme is to minimize with respect to u . This is a differentiable optimization problem and the solution is obtained by solving

$$(\mu A^T - \lambda \nabla)\mu^* = \mu A^T f + \lambda \Phi d \tag{8}$$

Next minimize (7) with respect to d . This optimization problem is element-wise decoupled. In the work, it is shown that the solution can be written explicitly as

$$d^* = shrink(\Phi u, 1/\lambda). \tag{i}$$

C. Split Bregman Methods for ROF

One of the simplest applications of the Split Bregman method is for ROF denoising. We give a brief overview of this technique here, and refer the reader to for more detail. In this image restoration model, the goal is to recover a denoised image, u , from a noisy image, f . The ROF model accomplishes this by solving an optimization problem of the form

$$\mu^* = arg \min_{\mu} |\nabla\mu| + \frac{\mu}{2} \|\mu - f\|^2 = arg \min_{\mu} |\nabla\mu|_1 + \frac{\mu}{2} \|\mu - f\|^2 \tag{9}$$

Note that this problem is of the form (5), when the operator A is taken to be the identity. To apply the Split Bregman approach to this optimization problem, we make the substitutions $d \leftarrow \nabla u$. Because this problem is defined on a two dimensional domain, $d = (dx, dy)$ is vector valued at each pixel. To approximately enforce these equality constraints, we add quadratic penalty functions. This results in the unconstrained problem

$$(\mu^*, d^*) = arg \min_{\mu, d} |d| + \frac{\mu}{2} \|\mu - f\|^2 + \frac{\lambda}{2} \|d - \nabla\mu\|^2 \tag{10}$$

We now wish to exactly enforce the equality constraints $d = \nabla u$. For this purpose, we apply Bregman iteration to the unconstrained problem (10). This results in the following sequence of optimization problems:

- 1: while 'not converged' do
- 2: $(\mu^k, d) = arg \min_{\mu, d} |d| + \frac{\mu}{2} \|\mu - f\|^2 + \frac{\lambda}{2} \|d - \nabla\mu - b^*\|^2$
- 3: $b^{k+1} = b^k + \nabla\mu^{k+1} - d^{k+1}$
- 4: end while

By putting all these pieces together, following very simple, yet efficient algorithm. Can be derived

D. Algorithm for Split Bregman ROF

- 1: while $\|\mu^{k+1} - \mu^k\| > \epsilon$ do
- 2: $\mu^{k+1} = GS_{ROF}(f, d^k, b^k)$
- 3: $d^{k+1} = shrink(\nabla\mu^k + b^k, \lambda)$
- 4: $b^{k+1} = b^k + \nabla\mu^{k+1} - d^{k+1}$
- 5: end while

E. The Split Bregman Method Applied to Globally Convex Segmentation

In this section, the Split Bregman method is applied to the GCS problem, to prove some elementary convergence results. As described in Algorithm 2, the convexified segmentation can be reduced to a sequence of problems of the form

$$\min_{0 \leq u \leq 1} |\nabla u|_g + \mu(u, r) \tag{11}$$

Where $r = (f - c_1)^2 - (f - c_2)^2$.

The problem is solved efficiently without the use of regularization. To this end, the Split Bregman method is applied. Just as it was done for the ROF model, introduce the auxiliary variable, $d \leftarrow \nabla u$, to weakly enforce the resulting equality constraint, then quadratic penalty function is added to get the following unconstrained problem

$$(\mu^*, d^*) = \arg \min_{u, d} |d| + \mu(u, r) + \frac{\lambda}{2} \|d - \nabla u\|^2. \tag{12}$$

In order to strictly enforce the constraint $d = \nabla u$, we apply Bregman iteration to the problem, just as was done for ROF. The resulting sequence of optimization problems is

$$(\mu^{k+1}, d^{k+1}) = \arg \min_{u, d} |d| + \mu(u, r) + \frac{\lambda}{2} \|d - \nabla u - b^k\|^2 \tag{13}$$

$$b^{k+1} = b^k + \nabla u^k - d^k \tag{14}$$

As described above, the problem (13) is solved using the alternating minimization scheme. It is initiated by considering the minimization of (13) with respect to u . As it is observed for the ROF problem, the algorithm converges very quickly even when an approximate solution is used. An approximate solution is derived using Gauss-Seidel method. This observation is in agreement with the previous literature on multiplier methods for differentiable problems. The use of inexact solvers in the context of differentiable problems is discussed in detail in [10], where conditions are presented under which multiplier methods can be guaranteed to converge [11].

E. Algorithm for Split Bregman for GCS

- 1: while $\|u^{k+1} - u^k\| > \epsilon$ do
- 2: Define $r^k = (c_1^k - f)^2 - (c_2^k - f)^2$
- 3: $u^{k+1} = GS_{gcs}(r^k, d^k, b^k)$
- 4: $d^{k+1} = Shrink_g(\nabla u^{k+1} + b^k, \lambda)$
- 5: $b^{k+1} = b^k + \nabla u^{k+1} - d^{k+1}$
- 6: find $\Omega^k = \{x: +u^k(x) > \mu\}$
- 7: update $c_1^{k+1} = \int_{\Omega^k} f dx$ & $c_2^{k+1} = \int_{\Omega^{k^c}} f dx$
- 8: end while

F. overlapped region and contour detection

As the chromosomes are spread all over the image. it has to be first identified in the image. Extract the information details with the shape of an image. Once the contour of the image is extracted the real characteristics of the image features will be identified. With the obtained image features the image is analyzed and classification of images is also done. Proper extraction of the contour produces more accurate features that help in image classification. The contour obtained may not have smooth boundaries or can have rough edges. To overcome this problem a polygon approximation can be performed. A polygon approximation is adopted which helps in smoothing the errors in the segmented boundary. The polygon approximation always requires a closed image boundary.

As the chromosomes are non-rigid in structure they can't be segmented accurately with already existing segmentation techniques. And the accuracy of the Segmentation result was only 90 % percent. Here to overcome this problem, Split Bergman active contour method is used for contour detection and to segment the chromosomes. As the Bergman active contour evolves based on the solution of elliptical curve, the curve accurately deforms according to the non-rigid structure of the chromosomes and thus increasing the accuracy of contour detection

Here with the help of same split Bergman active contour technique the overlapped region of chromosomes is detected in earlier works different algorithms were used for segmentation and for overlapped region detection

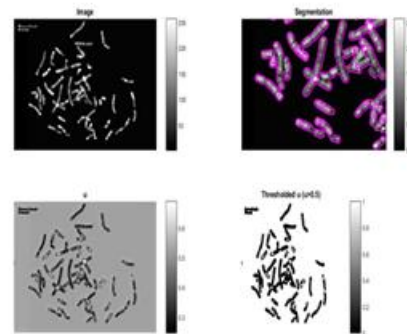


Fig-1 Overlapped region detected using Bergman Active Contour Technique

V. CONCLUSION AND FUTURE WORK

In this work, split Bergmen active contour method is used to segment and determine the overlapped region of chromosomes. Once the overlapped region is identified, it's easy to separate the chromosome and later it can be classified. This work makes use of single technique for both contour detection and overlapped region identification contrary to earlier works, where different techniques were adopted for contour detection and overlapped region identification.

In future work, it is planned to use Bergmen active contour method over M FISH chromosome images, evaluate its accuracy, to identify overlapped region and to separate and perform karyotyping.

REFERENCES

1. M. R. Speicher, S. G. Ballard, D. C. Ward, and Karyotyping, "Human chromosomes by combinatorial multi-fluor FISH," Nat. Genet., vol. 12, pp. 368-375, 2005.
2. E. Schrock, S. du Manoir, T. Veldman, B. Schoell, J. Wienberg, M.A. Ferguson-Smith, Y. Ning, D. H. Ledbetter, I. Bar-Am, D. Soenksen, Y. Garini, and T. Ried, "Multicolor spectral karyotyping of human chromosomes," Science, vol. 273, pp. 494-497, 2015.
3. T. Liehr and U. Claussen, "Multicolor-fish approaches for the characterization of human chromosomes in clinical genetics and tumor cytogenetics," Curr. Genom. vol. 3, pp. 213-235, 2014.
4. H. Choi, K. R. Castleman, and A. C. Bovik, "Joint segmentation and classification of M-FISH chromosome images," in Proc. 26th Annu.

- Int. Conf. IEEE Eng. Med. Biol. Soc., San Francisco, CA, Sep. 2004, pp. 1636–1639. Apr 2013
5. M. P. Sapat, A. C. Bovik, J. K. Aggarwal, and K. R. Castleman, “Supervised parametric and non-parametric classification of chromosome images,” *Pattern Recog.*, vol. 38, pp. 1209–1223, Aug. 2012.
 6. Y. Wang and K. R. Castleman, “Normalization of multicolor fluorescence in situ hybridization (M-FISH) images for improving color karyotyping,” *Cytometry*, vol. 64, pp. 101–109, Apr. 2011.
 7. W. C. Schwartzkopf, A. C. Bovik, and B. L. Evans, “Maximum-likelihood techniques for joint segmentation-classification of multispectral chromosome images,” *IEEE Trans. Med. Imag.*, vol. 24, no. 12, pp. 1593–1610, Dec. 2010.
 8. P. S. Karvelis, A. T. Tzallas, D. I. Fotiadis, and I. Georgiou, “A multichannel watershed-based segmentation .