Journal of Neonatal Surgery
ISSN(Online): 2226-0439
Vol. 14, Issue 11s (2025)
https://www.jneonatalsurg.com

Analysis of Chromosome Image Segmentation Using Image Processing Methods and DCE Antilogarithm

Dr. B. Vinothini[1], Ms. Maragatham T[2], Mr. S. Sivasubramaniam[3], Mr. V Suresh Kumar[4]

[1]Assistant Professor and Head, Department of Computer Applications, St.Joseph's College For Women, Tirupur

[2]Assistant professor, Builders Engineering College, Nathakkadaiyur, Tiruppur, Tamil Nadu

[3]Vice Principal, Assistant Professor in Computer Science, Terf's Academy College of Arts and Science, Tirupur, Tamil Nadu

[4]Head and Assistant Professor, Shree Venkateshwara Arts and Science College, Gobichettipalaym, Tamil Nadu,

.Cite this paper as: Dr. B. Vinothini, Ms. Maragatham T, Mr. S. Sivasubramaniam, Mr. V Suresh Kumar, (2025) Analysis of Chromosome Image Segmentation Using Image Processing Methods and DCE Antilogarithm. *Journal of Neonatal Surgery*, 14 (11s), 910-916.

ABSTRACT

Chromosome image segmentation is a crucial task in cytogenetics and genomics that enables accurate examination of chromosomal abnormalities and structures. In this study, we examine the effectiveness of the DCE (Dynamic Contrast Enhancement) Antilogarithm methodology for chromosomal picture segmentation when combined with image processing techniques. The DCE Antilogarithm can be used to contrast and enhance the visibility of chromosome regions that are often obscured by noise and difficult to identify in images obtained by microscopy. Image processing methods such as edge detection, thresholding, and morphological algorithms are used to distinguish individual chromosomes from complex, overlapping patterns. These methods are tested and validated on a range of chromosomal imaging datasets. Chromosome analysis and identification are essential for genetic research, and the proposed method aims to improve their accuracy and efficacy.

Keywords: *Chromosome, Image, Segmentation, Regions*

1. INTRODUCTION

Contour partitioning calls for an set of policies which could effectively find out the salient points notwithstanding the morphology of the chromosome. Curvature of the contour is one of the most normally used functions in the literature for detecting salient points that may be used for partitioning chromosomes. An important requirement is that the area of those salient factors want to be notably repeatable underneath various degree of object boundary noise. Curvature values of the chromosome boundary are incredibly susceptible to the boundary noise that is delivered via manner of exceptional staining and cell guidance techniques.

The DCE approach used for polygon evolution and form simplification for prunings purious branches in the centerline is proof towards boundary noise. This is because the noisy boundary points are deleted all through the preliminary degrees of the DCE algorithm mowing to the lack of contribution of those factors to the overall shape of the item. Furthermore the DCE approach additionally preserves the topological information of the particular item all through form simplification. These residences provide a notable platform for obtaining a fixed of initial salient factors at the contour of the chromosome outline which performs properly with obstacles regardless of their smoothness, yielding repeatable results. The ability to terminate the way of DCE form evolution ata given range of vertices further lends to its applicability. The requirement of the contour partitioning degree come to be to discover the 4 salient points of the chromosome which isolates the 2 telomere regions of the chromosome. However, because of morphological variations, the DCE give up result with sincerely four contour points (DCE rectangle) ought to not guarantee the inclusion of the factors required for setting apart the telomere place. Next it changed into empirically installed that a termination at 6 DCE factors could ensure that the required telomere give up factors could be retained in the set of candidate salient points. However, the 2 anchor factors (PE) of the skeleton acquired thru DCE based skeleton pruninga subset of these 6 DCE factors. Contour partitioning changed into completed via choosing the excellent 4 point combination that represents all the telomere give up elements.

The approach for selecting the best contour partitioning combination has twostages as listed below,

Training a feature based classifier using a large data set to capture desirable properties which contribute to make a good contour partitioning combination.

Using the trained classifier for selecting the best combination from a set of contour partitioning set of points

At the first stage, all the combinations across the data set were used as a pool of candidates to train the classifier. Use of this data aided the classifier to capture the desirable features across the data set and account for the high morphological variationsin human metaphase chromosomes. Once trained, the classifier was then usedto place the 12 candidates of each chromosome in the feature space separately. Thenthe signed Euclidean distance from the separating hyperplane was calculatedfor each of the candidate for the given chromosome. This distance was directly usedas a measure of the goodness of a given set of combinations. The best combinationfor partitioning the chromosome contour for isolating the telomere regionwas selected by picking the candidate with the largest distance from the separatinghyperplane. Unlike traditional rule based ranking algorithms, this approach requiredvery little high level knowledge of the desirable characteristics. The positioning of theseparating hyperplane encapsulated this high level information through user-providedground truth. Therefore the highest ranked combination was the best set of pointsthat could be used to isolate the telomere region to detect the evidence of sister chromatid separation.

Once the set of features were finalized, the combinations for the contour partitioningset of points were created for the complete data set. This created 16,800combinations for the total of 1400 metaphase chromosomes. Next, ground truthrequired for training and testing the accuracy of the classifier was collected. Theauthor examined all the possible combinations for each chromosome in the data setand manually marked the combinations which are viable solutions for partitioningand isolating the telomere regions of the chromosome. Next, a SVM classifier was trained and tested for effectiveness with 2 fold cross validation (50% - train data,50% - test data) and obtained an accuracy, sensitivity and specificity values of 94%,97% and 68% respectively. The results demonstrated the ability of the feature set to effectively detect good combination of candidate points for partitioning telomere regions. Slightly lower specificity suggests inclusion of some false positives into thedetection. However, this does not affect the accuracy of the contour partitioning sincethe algorithm picks the combination based on its rank as opposed to the classificationlabel.

Once the 12 combinations for a given chromosome was generated, the signeddistance from the separating hyperplane (Þ) was calculated. This distance is thegeometric margin of that sample in the feature space with respect to the separatinghyperplane. Then the combination of points with the largest positivegeometric margin (Þ) was selected and the telomere regions was segmented using thesepoints.

1.2 SHAPE EXTRACTION

The partitioned telomere curve segments carry important information relating to thepresence of premature sister chromatid separation. Traditionally, this is achieved byincorporating high level shape information using a large set of well defined featureswhich capture the subtle changes in shape characteristics. Once a large number of features have been defined, it is common to try and reduce the number of features tooptimize the feature space for classification. Principal component analysis (PCA) isone of the most common methods adopted to reduce the number of features. However,there are few disadvantages in a feature based set up. Namely,

- The performance of the system directly depends on the defined set of features. Therefore, this will lead to experimentation with a variety of feature combinations in order to optimize the setup.
- Some shape features are not invariant to the scale of the object. In these cases, additional computing is required to make the features scale invariant.
- The lack of expert knowledge of how each of the features contribute towards detecting the required shape.

1.3 NOVEL FUNCTION BASED ON ORTHOGONALITY

By definition, a set of functions is considered 'Orthogonal' with respect to a functional inner product < : ; : > within the domain [a; b] if,

$$< h_i, h_j > = \int_a^b h_i(t).h_j(t).w(t)\,dt \;=\; 0 \;\;, i \neq j$$

where, w(t) is the weight function defined for the same domain. A given function f(t)is expressed as a linear combination of coefficients calculated related to an orthogonal basis.

$$f(t) \;=\; \sum_{i=0}^{\infty} \alpha_i.h_i(t)$$

Where each coefficient $\alpha_i$can be calculated as follows,

$$\alpha_i = \frac{< f, h_i >}{\|h_i\|^2}$$

Instead, we need to calculate moments of different orders using these coordinatevalues (for x and y coordinates separately). By definition, the kth order moment ofthe x coordinate is given by,

$$\mu_k = \int_a^b x^k f(x) dx$$

The coordinate points in the partitioned contour segments are not uniformlydistributed. This is due to the forces acting upon each control point in the GVFsegmentation stage. Therefore, each contour segment has to be parameterized accordingly.

$$\int_i^{i+1} \lambda^k f(\lambda) d\lambda = \frac{d1(i+1)^{k+1} - d(i)^{k+1}}{k+1} \times \frac{f(i+1)+f(i)}{2}$$

Therefore it is assumed that calculation of ten moments accurately recreate orrepresent a shape characteristics at the telomere region. This is a reasonable thresholdsince the inclusion of moments of higher order values adds very little informationof the general shape while having the possibility of including boundary noise of thechromosome outline.

$$P_n(t) = \frac{1}{2^n n!} \frac{d^n}{dt^n} (t^2 - 1)^n$$

k is the maximumorder of the moments and L is the total length of the curve segment.

$$\hat{\alpha}_k = (-1)^k \frac{2k+1}{L} \times \sum_{i=0}^{k} (\frac{-1}{L})^i \binom{k}{i} \binom{k+i}{i} \mu_i$$

$$\hat{\hat{\alpha}}_i = \hat{\alpha}_i \frac{b-a}{f_{max} - f_{min}} + \delta_{i,0} \frac{a.f_{max} - b.f_{min}}{f_{max} - f_{min}}$$

With the imposed limitation of 10 coefficients per axis, this yielded 20 featuresrepresenting shape features along both x and y axis. A second SVM classifier wastrained using 90 labeled set of telomere coordinate curve segments to effectively detectshape variations inherent to telomere regions with sister chromatid separation.A multi-layer perceptron kernel was used with Quadratic programming was was utilized for training this classifier. Thesetelomere curve segments were examined for evidence of sister chromatid separationwhich appears as a boundary concavity and was manually labeled for training andtesting the classifier. The support vector machine classifier is a 'large margin classifier', which maximizes the largest distance to the nearest training data points ofany class. This yielded more reliable classification for any new data points evenclose to the decision boundary/plane. With 2-fold cross-validation, the support classifier demonstrated accuracy higher than 92%.If a telomere region was detected for sister chromatid separation, then the end pointcorrection discussed and altered so that the extended line satisfy thecoordinates of the telomere midpoint. This correction was not meantfor correcting the centerline of the chromosome for the artifact of sister chromatidseparation. Instead it attempts to split the contour of the chromosome into two approximatelysymmetrical contour segments which is a requirement for the Laplacianbased thickness measurement algorithm proposed.

1.4 Image Registration

The misalignment of spectral images is an inevitable phenomenon dueto the fundamental optical properties of the microscopic imaging system.When it occurs, it adversely affects the classification accuracy on pixels, especiallyon the edges of chromosomes. However we found only four imagesin the database that exhibited a noticeable misalignment, which are not fromthe optical properties but from some other unknown source of errors. Whileimage registration of multichannel or multitemporal images is an active researcharea, we do not find a need for an image registration algorithmfor M-FISH images. Therefore in this section we will briefly discuss the fundamental cause of misalignment of M-FISH images, and present an exampleof misaligned M-FISH images and its effect of classification accuracy due tomisalignment.

1.5 Source of Misalignment

A basic principle of optics states that the focal length changes dependingon wavelength. Thus when wavelength changes, the in-focus plane ofan object changes, resulting in axial chromatic aberration. Also the magnificationis inversely proportional to the focal length. Thus depending on thewavelength, the magnification also changes, which results in lateral chromaticaberration. Both aberrations can be found, even in the best currently availableobjectives. Therefore, when multiple emission wavelengths are used to imagethe same object, as in fluorescence imaging, chromatic aberration is inevitable.Furthermore, mechanical vibration induced by filter changes may cause misalignment.Classification accuracy on pixels near the edges of chromosomesare affected when misalignment occurs.

Misalignment due to chromatic aberration should be consistent throughoutall images. Variations in the amount of misalignment should be due toother factors such as any mechanical vibrations. Three channels are displayed as a colorimage: (DAPI in blue, Gold in green, and Far red in red channel). DAPI andFar red are selected since they have the farthest distance in wavelength, andthe corner of an image is shown since the lateral chromatic aberration becomes more severe as the distance gets further from the optical axis. As the figureshows, the amount of misalignment is negligible (even not noticeable). In fact,all the images we have observed in the database had a negligible amount ofmisalignment. The misclassifications on the edges of chromosomes commonlyoccur because intensities on those pixels are weak, making less certain of theirmemberships, and also when a chromosome appears larger due to bloomingon one channel than on another channel, the non-overlapping area around theedge of the chromosome is misclassified.

1.6 GENETIC APPROACH

Genetic Algorithms (GA) are stochastic search procedures introduced by J.Holland in the 70's. These algorithms are based on ideas and techniques from genetic and evolutionary theory which is a field of artificial intelligence and is a kind of searching for good solutions that mimics the process of natural evolution. GAs generate valuable solutions for hard optimization problems using techniques that are inspired by natural evolutionary operators, such as inheritance, mutation, selection, and crossover.

1.6.1 GA’s Population

In GAs, there is a population containing a number of solutions which are represented by strings (called chromosomes or the genotype) that evolve in the direction of better solutions. Each string is an encoded candidate solution. Conventionally, solutions are encoded in binary strings of 0s and 1s, but other kinds of encoding models are also probable. The evolution starts by generating several individuals to create an initial population. Then, the fitness function is computed for each individual to produce a selection priority for individuals throughout the generations. Therefore, individuals are preferred from the present population according to their fitness values and modified to a number of offspring. The new population replaces the current population and is used as an input to the next iteration of the algorithm. Usually, the algorithm will be terminated when either maximum number of generations is reached, or a reasonable fitness value is attained.

1.6.2 GA Criteria

A common genetic algorithm involves two main parts:

1. All solutions should have a genetic representation (in a shape of chromosome)

2. There should be a fitness function to assess the solutions,

An array of binary bits is a typical representation of a solution. However, other types of representation can also be used.

1.7 Fitness Function

Regarding optimization problems, to produce better solutions from the current one, there should be a fitness function to evaluate the quality of each encoded solution through the generations. The form of the fitness function is always a problem-dependent issue. Fitness value is a quality measurement of each solution. Better fitness values belong to better individuals in each population. When termination criteria are satisfied, algorithm reaches to better fitness value. In the final generation, a solution with better fitness value among others is found as the desired solution. As an example of computing the fitness value, knapsack problems are explained here. The best solution for these problems is a representation that maximizes the total value of the objects that can be located in a knapsack with a limited capacity. This solution might be illustrated by a binary array of bits. Each object is represented by one bit, which has the value of 0 or 1. Value of 0 means object is not selected and 1 means object is selected to put in the knapsack. Putting objects in knapsack continues until the capacity of knapsack becomes full or it would not be enough for next object. As a result, several random individuals are generated. The fitness of each solution is the sum of values of objects that have a value of 1 in their related bit of the binary array or the chromosome.

1.7.1 Initialization of the Population

For the purpose of creating an initial population, many individuals (solutions) are generated randomly. Usually, the population size involves hundreds or thousands of probable solutions. Subsequent to generate first population and define fitness function, a new population is produced through the mutation, crossover, inversion and selection operators.

1.7.2 Selection Procedure

An important procedure during GA generations is how to select individuals from the current population and use them to generate next population. Selection operation directly depends on the individual quality. The quality of each individual is measured by a fitness function. Individuals with better fitness value are more likely to be selected. There are several selection algorithms that can prefer better individuals from a population such as roulette wheel selection and tournament selection operators.

1.7.3 Reproduction Process

Reproduction methods are biologically inspired and based on the use of two parents to produce a child. To produce each new child, two individuals (parents) are selected from the current generation to reproduce new individuals for next generation through the crossover operation and/or mutation. Inherently, a generated child shares many characteristics of its parent through the swapping of some genes between parents. Therefore, the new child takes advantages from both parents and might be better in comparison to the previous individuals. Reproduction process will be terminated when the size of the new population is equal to the size of the last population.

1.7.4 Application to Image Segmentation

Recently, a variety of evolutionary optimization algorithms has been introduced in order to optimize the image segmentation results. In order to achieve this aim, genetic algorithms which are vital components of the wider class of evolutionary algorithms are utilized. Therefore, an image segmentation method with high accuracy and localization ability is generated. In this method, the initial segmented image is represented by means of a chromosome, where each gene of the chromosome indicates a separated region of the image. In order to generate the primary population, several chromosomes will be produced based on the initial chromosome. The other populations will be generated accordingly.

1.8 Chapter Summary

The proposed method wasobserved to perform satisfactorily despite the high morphological variations on cellsimages from DAPI as well as Giemsa stained images. Furthermore,the proposed method performed better at calculating width profile of chromosomeswith minimal influence from boundary noise than the centerline based approach. Dicentric chromosomes appear in low frequencies in human metaphase cell imageseven at considerable radiation levels and become even less frequent in lowerradiation dosages. Therefore, it is paramount to include all types of chromosomes inthe analysis for dicentric detection. This is a major drawback in methods currentlyknown. The candidate based approach in the proposed algorithm gives theability to include both acrocentric and submetacentric chromosomes into the analysis.Coupled with the CBCC (Candidate Based Centromere Confidence) metric, theproposed algorithm is geared to provide useful information to the expert involvedin the diagnosis process. Its important to notice that though these are essential requirementsfor radiation dosimetry, they are also desirable properties to have in anychromosome analysis and centromere detection algorithm.More research is warranted in extending the centromere detection algorithmto accurately identify dicentric chromosomes.

Two new unsupervised nonparametricclassification methods for M-FISH images were presented, which are a fuzzylogic classifier and a template matching algorithm. Both methods providea significant advantage in terms of computation time compared to supervisedmethods, and their accuracies were comparable to that of a maximumlikelihoodclassifier.Thus the both geometry and pixel classification results have to be merged inorder to achieve better segmentation results.In this chapter, we present a novel decomposition method for overlappingand touching chromosomes that utilizes the geometry of a cluster, pixelclassification results, and chromosome sizes. We also introduce basic elementsof overlap and touching cases. These basic elements yield hypotheses of possibleoverlapping and/or touching cases. Given a cluster, multiple hypothesesare evaluated, and the most likely hypothesis is chosen as the correct decomposition.

Various pixel classification methods for M-FISH imagesare described, which include supervised parametric, supervised nonparametric,and unsupervised nonparametric methods including two new classificationmethods for M-FISH images that do not require training of a classifier (unsupervised)nor require class parameter estimation (nonparametric).Given a number of objects, the choice of classifier depends on the knowledgeabout the samples in the feature domain such as the number of classes,the prior probabilities, the forms for the class-conditional probability densityfunctions, the values for the density functions, and the category labels of thesamples.When the labels are available, we can learn the statistical propertiesof the samples and design a classifier that utilizes that knowledge. Overlapping and touching chromosomes were effectivelydecomposed using the developed decomposition method. Given a cluster, anumber of hypotheses were formed utilizing the geometry of a cluster, pixelclassification results, and chromosome sizes, and a hypothesis that maximizedthe likelihood function was chosen as the correct decomposition. After chromosomes are individually identified, misclassified pixels were effectively correctedwhile preserving the trans-located pixels, using the prior adjusted reclassificationmethod.

Image segmentation is normally used to detect different objects and their boundaries in images. In particular, image segmentation is the procedure of conveying a label to each pixel such that pixels with similar labels share firm visual characteristics. As it is mentioned previously, there are several segmentation methods, such as Edge-based, Region-based, Clustering-based methods and Split/merge approaches. Brief descriptions of these approaches are given in subsequent sections. GOIS can simultaneously gain the following five objectives: (1) detected edges are always continuous; (2) final detected regions can indicate significant objects from the image; (3) short computation time; (4) this method is practical with acceptable performance for noisy images; (5) high accuracy in comparison with other segmentation methods. Elevated performance of GOIS is shown in terms of assessment performance and computation time compared to some well-known segmentation methods.

2. CONCLUSION

In this work, the DCE Antilogarithm methodology for chromosomal picture segmentation is used with image processing techniques. This combined strategy makes it possible to identify and segment individual chromosomes more precisely by increasing contrast and decreasing noise. The results demonstrate notable gains in segmentation ability as compared to conventional techniques since chromosomal structures are hidden by overlapping or low-quality images. Chromosome analysis can be successfully accomplished by combining edge detection, thresholding, and morphological approaches with the improved contrast provided by the DCE Antilogarithm. Better diagnostic tools and a deeper comprehension of genetic abnormalities are made possible by the potential for wider genomics applications of this technology. Its feasibility for use with more complicated data and its integration with automated analytical tools for thorough genetic studies may be further investigated.

REFERENCES

[1] N. Xie, X. Li, K. Li, Y. Yang and H. T. Shen, ”Statistical KaryotypeAnalysis Using CNN and Geometric Optimization,” in IEEE Access,vol. 7, pp. 179445-179453, 2019, doi: 10.1109/ACCESS.2019.2951723.

[2] [M. S. Al-Kharraz, L. A. Elrefaei and M. A. Fadel, ”Automated Systemfor Chromosome Karyotyping to Recognize the Most Common Numer-ical Abnormalities Using Deep Learning,” in IEEE Access, vol. 8, pp.157727-157747, 2020, doi: 10.1109/ACCESS.2020.3019937.

[3] H. M. Saleh, N. H. Saad, N. A. Mat Isa, ”Overlapping ChromosomeSegmentation Using UNET: Convolutional Networks with Test TimeAugmentation ” in Procedia Computer Science, Vol 159, pp. 524-533,2019.

[4] H. A. Al-Ameri and W. Al-Hameed, ”New algorithm for separationoverlapping & touching chromosomes”, 2020 J. Phys.: Conf. Ser. 1530.

[5] E. Poletti, E. Grisan, A. Ruggeri, ”Automatic classification of chromo-somes in Q-band images” in Annual International Conference of theIEEE Engineering in Medicine and Biology Society, 2008, 1911–1914.

[6] J Liang. Decomposition of overlapping chromosomes. In: C Lundsteen and J Piper (ed). Automation of Cytogenetics, Springer, Berlin, 1989b, p. 177-90.

[7] J Liang. Fully automatic chromosome segmentation. Cytometry 1994; 17, 196-208.

[8] G Agam and I Dinstein. Geometric separation of partially overlapping nonrigid objects applied to automatic chromosome classification. IEEE Trans. Patt. Anal. Mach. Intell. 1997; 19, 1212-22. [10] B Lerner, H Guterman and I Dinstein. A classificationdriven partially occluded object segmentation (CPOOS) method with application to chromosome analysis. IEEE Trans. Sig. Proc. 1998; 46, 2841-47.

[9] Prabhu S Abnormal Chromosome Identification Using Hybrid Image Segmentation Method , International Journal of Advanced Science and Technology , Volume XXIX, Issue VIII, January/2020 **0.41**.

[10] M Popescu, P Gader, J Keller, C Klein, J Stanley and C Caldwell. Automatic karyotyping of metaphase cells with overlapping chromosomes. Comp. Biol. Med. 1999; 29, 61-82.

[11] GC Charters and J Graham. Trainable grey level models for disentangling overlapping chromosomes. Patt. Recog. 1999; 32, 1335-49.

[12] GC Charters and J Graham. J. Disentangling chromosome overlaps by combining trainable shape models with classification evidence. IEEE Trans. Sig. Proc. 2002; 50, 2080-85.

[13] R Ogniewicz and M Ilg. Voronoi skeletons: Theory and applications. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Illinois. 1992, p. 63-9.

[14] M Styner, G Gerig, J Lieberman, D Jones and D Weinberger. Statistical shape analysis of neuroanatomical structures based on medial models. Med. Image Anal. 2003; 7, 207-20.

[15] CM Boesse, MK Henry, TW Hyde and LS Matthews. Digital imaging and analysis of dusty plasmas. Adv. Space Res. 2004; 34, 2374-78.

[16] PK Sahoo, S Soltani and AKC Wong. A survey of thresholding techniques. Comp. Vis. Graph. Image Proc. 1988; 41, 233-60.

[17] RM Haralick and LG Shapiro. Computer and Robot Vision, Vol. I, Addison-Wesley, Reading, MA, 1992, p. 1-608.

[18] N Otsu. A threshold selection method from gray-level histograms. IEEE Trans. Syst. Man Cybern. 1979;.

[19] Prabhu SProposed Hybrid Segmentation Algorithm Using Preprocessing, Shape Extraction, Source Of Mis-Alignment And Genetic Approach In Digital Image Processing, The International journal of analytical and experimental modal analysis , Volume XI, Issue XII, December/2019.