ISSN: 0974-2115
www.jchps.com Journal of Chemical and Pharmaceutical Sciences

A Review of Human Chromosome Disentanglement Strategies

Mary Sylvia Jeny Priya*, Helen Sulochana
Dept. of Electronics and Communication Engineering, Ponjesly College of Engineering, Nagercoil, Tamilnadu.
***Corresponding author: E-Mail: sylvia.jeny.16@gmail.com**

ABSTRACT

In cytogenetics, automated chromosome analysis is an important task which involves segmentation of chromosomes and their classification into one of 24 groups. An effective automatic analysis would greatly help the cytogeneticist usual work. The important step in the analysis of chromosomes is the segmentation step. In this paper, we analyzed eight disentangling methods and compared their performance in segmenting chromosomes and chromosomes clusters. A detailed analysis of the results highlights that, even though every single algorithm shows their own strong/weak points, the method using geometric evidence and image information has the overall best performance.

KEY WORDS: Karyotype, chromosome disentanglement, variants, pale path, single - chromosome likelihood.

1. INTRODUCTION

Chromosome karyotyping is an essential task in clinical and cancer cytogenetic laboratories for screening and diagnosis. Chromosome karyotyping is the process of arranging all the chromosomes in a cell graphically, according to an international system for cytogenetic nomenclature (ISCN) classification. Chromosomes are first need to be stained with fluorescent dye and then it is imaged through a microscope for successive analysis and classification. Each chromosome in the image has to be identified and assigned to one of 24 classes: the result so called karyotype image. The most common methods of dye-based chromosome banding are G-(Giemsa), R-(reverse), C-(centromere) and Q-(quinacrine) banding. Q- bands are considered equivalent to G- bands. G- and R- bands are the most commonly used techniques for chromosome karyotyping. As the banding patterns of metaphase images are clearly visible, most of the studies concentrates on such metaphase images. Figure 1 shows the four different bandings of chromosome images.

The first step in chromosome analysis is segmentation, which includes resolving touching and overlapping chromosomes. There is a variety of disentangling methods proposed in the literature for both banded and unbanded chromosome images. All these methods are based on five main ideas: skeleton, shape concavities, shape validation, pale paths (between touching chromosomes) and classification driven segmentation. Most of the authors use any one of these techniques and some many combine one or more techniques for disentangling the chromosomes.

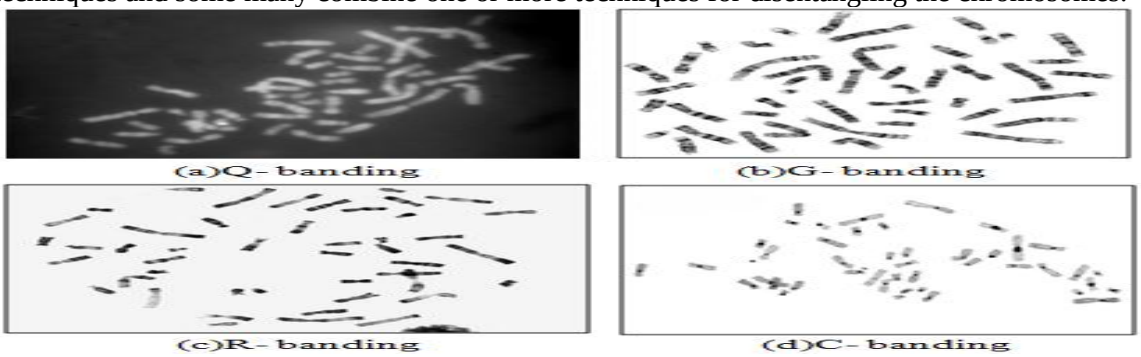

Fig.1.Different bandings of chromosome images

Geometric descriptors such as concavities and boundary curvature along the object boundaries are used for detecting candidates for cut points. Concavities are usually detected by points of large negative deflecting angle or large negative curvature or by local methods. In both convex and concave points are used for separating the cluster, where concave points are used to define the cut points and convex points are used to define the end points. Another important descriptor is the skeleton, which is a 1D representation of components, by strokes of thickness one in its interior. The nodes of the skeleton indicate touching and overlapping points of the chromosomes. The boundary in their neighborhood is searched for finding the potential cut points. Some works uses pale paths between concave points as potential cut points.

Usually more cut points are obtained than actually needed. From all the cut points, some suitable cut points need to be selected. Some of the authors validate the potential cuts, by using geometric information about the components. In classification-based method the cut points are selected after classifying the corresponding chromosome segments. However, this method is used by the authors for separating clusters of two chromosomes only.

Ji (1989), uses 'pale path' technique, Agam and Dinstein (1997), uses boundary curvature to separate touching and slightly overlapping chromosomes, Graham (2002), tried to integrate the segmentation and classification steps to resolve overlaps and adjacencies. Lerner (1998), proposed a classification driven segmentation.

Ritter and Gao (2008), proposed a method that relies on the generation of number of possible chromosome configurations called 'variants'. Grisan (2009), uses space variant thresholding scheme and introduces single chromosome-likelihood (SCL) for cluster analysis. Karvelis (2005), uses watershed transform and gradient paths; moreover this can be extended for multi-spectral images.

Even though these algorithms give better results, they have some limitations. The aim of this work is to measure and compare the performance of all the above quoted algorithms. The performance is analysed based on accuracy, the type of images they have used for testing their algorithms and the ability of their algorithms in resolving complex images.

The layout of this paper is as follows: In Section II, the data sets used by the authors considered in this work, is given. Section III gives a brief explanation about the eight disentangling methods. In Section IV, the segmentation results of these eight methods are shown and in Section V those methods are discussed. The conclusion of the paper is presented in Section VI.

Data Sets: Most of the authors considered in this work, uses their own database of images for testing their algorithm, and few authors uses publicly available database. Ji (1989), uses his own database for testing, in which the cells are stained homogeneously with Giemsa or Orcein and digitised automatically on the MRC's Fast Interval Processor (FIP). 256 homogeneously Giemsa stained G-banded cells and 457 Orcein stained unbanded cells from the data set were selected for testing. Ji (1989), uses the dataset which comprised of 458 clusters of G-banded chromosomes selected manually from 125 cells and 565 clusters of Orcein stained unbanded chromosomes selected visually from 102 cells. The overlapping pairs present in the database were ignored during testing. Karvelis (2005), uses database of images collected at the Department of General Biology, Medical School of University of Ioannina. The database consists of 940 chromosomes: 515 isolated, 396 touching and 29 overlapping. Popescu (1999), created their own chromosome image library which consists of 3979 chromosomes that includes isolated, non-overlapped, non-rotated chromosome images extracted from 209 cell images. The training set consisted of 3582 chromosomes, approximately 149 chromosomes per class, with the test set containing 397 chromosomes, approximately 16 chromosomes per class. Ritter and Gao (2008), forms Pki-3 dataset from the already available dataset Pki. The Pki dataset consists of 971 Giemsa stained pro- and amnion and blood cells at pro- or metaphase compiled at the local cytogenetical institute in Passau/Germany for routine cytogenetic screening. All cells at late metaphase, characterized by split arms, were removed, and a new dataset was created with the remaining cells. The new dataset thus formed was named as Pki-3 and it contains 612 cells with 28148 chromosomes. Lerner (1998), uses database of manually collected images, which consists of around 300 images with approximately 100 from each class of 'brokens' and 'wholes'. This dataset consists of around 481 chromosome images. The dataset used by Grisan (2009), is composed of 162 images, containing a total number of 6683 Q-banded chromosomes. Graham (2002), uses a publicly available '600-band' dataset containing 6177 chromosomes obtained from 136 G-banded cells. For each chromosome, there is an isolated image, the class specified by a cytogeneticist, and an identifier for its cell of origin.

2. METHODS

In this section, we briefly describe all the disentangling techniques we have taken into account. We analyzed each methods and performance is compared based on their accuracy in resolving touching and overlapping configurations of chromosomes.

Boundary Curvature and Pale Path: In 'boundary curvature and pale path' technique, Ji (1989), describes the method for splitting the chromosomes in a cluster with only few chromosomes. The major characteristic of this method is determining a cut point and relevant associated path. For this, they proposed two algorithms. In the first algorithm, first cut point is taken as the reference point. The distance between the reference point and every cut point on the boundary is calculated. Then local minima of this distance function are found and the one with least distance value is selected. The vertex corresponding to these local minima is taken as a candidate for the second cut point. If the distance between these two cut points is less than a threshold, then that is the required path. If the distance between these two possible cut points exceeds the threshold then they proceed with the second algorithm.

Algorithm 2 generates a path beginning at the first cut point based on the direction, search range and criterion of density. Initial direction of the path is the bisector of the boundary angle at the first cut point. The direction is recomputed after every'd' pixels. Search range is the set of points lying within the object whose displacement from the current point is not more than 3 pixels in the search direction, nor more than 3 pixels in the perpendicular direction. At each step, the least density points are found in the two directions of the search range. The densities of the pixels which lie along the direction are weighted and the one which has least density is chosen as the 'next' point in the path. The search then starts again from the next point and continues until the boundary is reached. In [28], authors uses this boundary curvature and pale path technique in their work, for detecting cut points and cut lines, to separate the overlapped chromosomes.

CPOOS Method: The idea behind Classification-Driven Partially Occluded Object Segmentation (CPOOS) method is, classification of segmented images created by connecting points of high concavity along curvatures, to resolve

partial occlusion in images. Initially, isolated objects in the image are classified into their classes. Thereafter the CPOOS method is applied to those clusters in the image that could not be separated. Clusters of touching objects are identified by their size and their failure to be classified as one of the object classes. To draw possible separating lines between touching objects, 'r' most concave points along the curvature are selected and considered as potential cut points. From those 'r' potential cut points, the points which are closer than a distance 's' from more acutely concave points are eliminated. The values used for 'r' and's' are chosen as 8 and 10 respectively. Finally, separation lines are obtained by connecting the pairs of cut points. The two segments formed by the separation line are assumed as touching objects. A trained classifier with 'apriori' knowledge of the identity of two chromosomes that compose the cluster is then used to verify the corresponding hypothesis. A multilayer perceptron (MLP) trained by back propagation learning algorithm is used to select correct separation line.

Recursive Watershed Transform: Initial segmentation of the image is created using the watershed transform. Next, watershed transform is recursively applied to every segmented area, until there are no new areas. The final segmented chromosome areas are formed by computing all paths starting from high concavity points. For a measure of separation of points in the image, the distance transform is used. The number of regional minima of the negative distance transforms acts as an indicator to the watershed transform which depicts the number of objects that need to be segmented.

The distance metrics used to compute the distance transform include

$$L_1: |x_1 - x_2| + |y_1 - y_2| \quad \text{... (1)}$$

$$L_2: \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad \text{... (2)}$$

$$L_\infty: \max(|x_1 - x_2|, |y_1 - y_2|) \quad \text{... (3)}$$

When applying watershed transform, over segmentation results due to the fact that chromosomes are not completely circular to present exactly one minimum to the distance transform. Hence, grayscale reconstruction is used to reduce the number of maxima-minima in a grayscale image.

The grayscale reconstruction $p_I(J)$ of (mask) I from (marker) J is:

$$\forall p \in D_b \quad \text{... (4)}$$

$$\rho_I(J) = \{k \in [0, N-1] \mid p \in \rho_{T_k}(I)(T_k(J))\} \quad \text{... (5)}$$

$$T_k(I) = \{p \in D_I \mid I(p) \geq k\} \quad \text{... (6)}$$

Where I, J be the grayscale images and p be the pixel. The grayscale transform is applied to reduce the number of maxima of the distance transform as follows:

$$D_h(D) = \rho_D(D - h) \quad \text{... (7)}$$

Where D the distance transform of the binary image, D_h the grayscale reconstructed distance transform. For separating very difficult cases of touching chromosomes, pale path technique is used.

Variant Analysis: Variant analysis technique consists of combination of two phases, rule-based phase and phase driven by constraint discriminant analysis. In first phase, prototypical shape elements related to touching and overlaps are identified recursively and in the second phase, complete and ambiguous cases are treated by using global context and variant analysis.

Initially, shape descriptors such as skeleton, polygon approximation and boundary curvature are computed from all components. The information related to skeleton such as crossings, branching and their Euclidean distances from the complement is collected and represented in skeleton groups. The clear phase detects overlaps, bridges, M's and T's and cuts them. Then, 'connected components-shape descriptors-segmentation' is re-iterated until no more cuts are executed.

The next ambiguous phase recursively creates local variants (feature set polarities and shapes) and organise the variants in a data structure. All global variants with an appropriate number of components are extracted from the data structure. These competing segmentations were classified with the constrained classifier. The solution with the best score was retained.

Hypothesis and Test: This method combines classification evidence and shape evidence based on trainable shape models. The approach adopted here is called 'hypothesis and test'. Based on some data driven process, candidate objects are initially hypothesized and the hypothesized objects are then evaluated against a model of their expected appearance. The model of appearance of the objects in this case is long and flexible with a recognisable banding pattern. Classification evidence method uses short sequences of band pattern that are not concealed by the overlapped region. Each segment is matched with large set of trained templates (partial chromosome models (PCMs)) to yield a feature vector for linear classifier. For shape evidence, a method for representing shape using trainable models was used. They used 'Chord distribution model' (CDM) for shape modelling.

Geometric Evidence and Image Information: In this paper, touching and overlapping chromosomes are recursively resolved by searching a tree of choices. Based on image information and geometric evidence, the best combination of cuts and overlaps are chosen. Here segmentation and disentanglement is done without any

assumption on their number or dimension distribution. To assess whether an object is a single chromosome or a cluster, and to score each object, it describes a measure of single- chromosome likelihood (SCL).

Given a blob C, with area A_c and with the description of its axis given by the curve $a(l) = [x(l), y(l)]$ onto which the curvature $k_a(l)$ is computed, the score of C is

$$SCL(C) = \begin{cases} 0 & , \ if\ bulges \\ 9 - \int k_a(l)dl, & if\ A_C < th_A \\ 10 - \int k_a(l)dl, & otherwise \end{cases} \quad \ldots (8)$$

Where th_A is the area threshold that was heuristically set at 150 pixels and $k_a(l)$ is the curvature of contour 'a' and it is computed as follows:

$$k_a(l) = \frac{\dot{x}(l)\ddot{y}(l) - \ddot{x}(l)\dot{y}(l)}{(\dot{x}^2(l) + \dot{y}^2(l))^{\frac{3}{2}}} \quad \ldots (9)$$

The object that does not pass the SCL test becomes the root of a tree. Each node represents one of several cuts and overlaps hypotheses, which is generated based on image and geometrical evidence. The final disentanglement is obtained by recursively analysing each node, until the leaf of the tree is evaluated. Authors use length and area of split chromosome images for evaluating the best combination of hypothesis.

Valley Searching: In valley searching algorithm, a touching cluster should be split by starting from a boundary concavity and following a path having lower density value than the average density of the chromosomes. The essence of the procedure is to find a 'cut point' at a boundary concavity and a path to the opposite boundary, which passes through a region with relatively low density using a 'valley following' technique. By deleting points along this path and redefining connected components, two separate objects are obtained. It is based on two hypotheses: where chromosomes touch, the boundary of the cluster tends to form concavity. The density in the touching region, particularly where it meets the cluster boundary, is relatively low. This method does not assume that a path passing through the touching region is straight. The initial direction of the path is the bisector of the boundary angle at the cut point. Direction is recomputed after every five points of the path have been found. The pixels ahead of the path and on either side are searched for the minimum density pixel which is chosen as the next path point provided that there is no rapid direction change. The path ends when it reaches the opposite boundary.

Automatic Karyotype System (AKS): AKS has four novel features. The first procedure is for finding medial axis using cross section sequence graphs (CSSG). The second feature uses a 'pale path' concept for splitting banded touching chromosomes. The third feature, extracts a set of primitives from a chromosome cluster. The final feature is the use of mathematical programming models to find optimal assemblies of primitives.

3. RESULTS AND DISCUSSION

The segmented results of the eight methods considered in this work are shown in the following figures. Ji's 'boundary curvature and pale path' technique is simple and fast. It works well both on banded and unbanded chromosomes. A single cut point is needed to find a path. Most of the parameters can be adaptively changed according to an initial result. The problem found in this technique is, if two chromosomes are tightly packed together but does not overlap completely; there will not be any 'pale' path for separation. Sometimes the result is false positive, with the chromosomes being split at the centromere. If we know how deep the chromosomes are tightly packed or if the length of the inserted chromosome is known, we can then identify the correct split. But the length of chromosome is unknown until it is known that which class it belongs to.

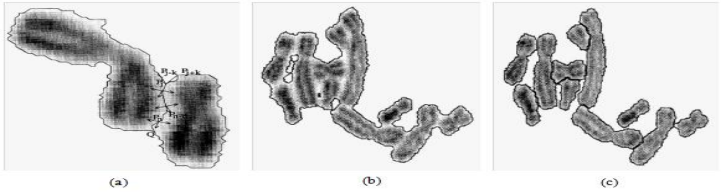

Fig.2. Ji's approach for chromosome image segmentation (a) direction of pale path(b) original unsegmented image(c)final segmented image

Figure 2 shows the pale path concept of Ji (1989). The figure 2(a) shows the search range, directions, angle and resultant path. Searching starts at point P_j. The first direction is indicted by angle P_{j-k}, P_j, P_{j+k}. Small arrows indicate the following search ranges and directions. There is a big boundary curvature at P_j. The point P_j is the possible cut point, which is confirmed by the pale path. Fig 2(b) shows the original image and fig 2(c) shows the resultant segmented image obtained by using Ji's method (1989).

Next, the variant analysis technique proposed by Ritter and Gao (2008), is based on global statistical information on karyogram to select the correct cuts from the proposed ones. It consists of two phases. In the first phase, shape elements indicating touching or overlaps are detected. In the second phase, ambiguous cases are treated. Prototypical shape elements are detected in both phases, in a conservative or in an offensive way.

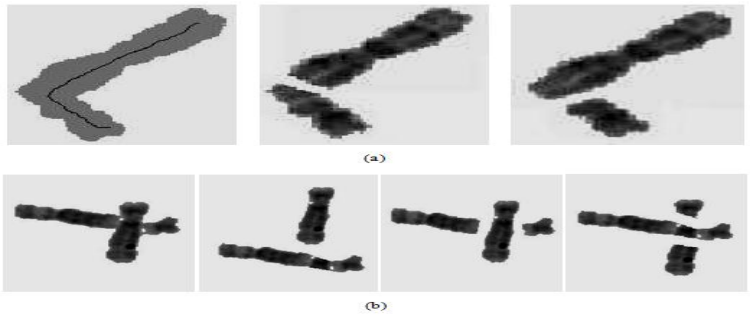

Figure.3.Segmentation results of variant analysis technique (a) an L-configuration with its skeleton and its two variants (b) an ambiguous X-configuration and its three variants

Relying on the clear phase leads to complex rules and descriptors; similarly, relying on the ambiguous phase alone leads to a combinatorial explosion in majority of cases. Therefore, they try for the balance of the two. The problem arises in this method are: bent chromosomes are hard to differentiate from two chromosomes touching at their tips; narrow centromeres resembling two lightly touching chromosomes in a line; artificial holes in chromosomes looking like genuine holes between touching and overlapping chromosomes. Also to describe the shape elements about 170 parameters are needed and also it requires runtime of few hours for complex cases. Figure 4 shows the result of segmentation obtained by using the variant analysis technique. Figure 3(a) shows an L-shaped configuration and its variants, and after classification it is found that the second variant is correct. Similarly figure 3(b) shows an ambiguous X-shaped configuration and its three variants, and the third variant is found to be correct.

Figure 4 shows the result of chromosome splitting step obtained by using the algorithm proposed by Popescu (1999). The cross section sequence graphs (CSSG) proposed by Popescu (1999), is less sensitive to noise and produces more perceptive junctions than standard skeletonization method. For splitting banded chromosomes, this system uses pale path concept. Figure 5 shows the segmentation result obtained by Karvelis (2005), which is based on recursive watershed transform. Recursive watershed transform, recursively applies watershed transform for the separation of touching objects in images. But this work resolves only touching group of chromosomes and not overlapping chromosomes.

Fig.4.Result of the chromosome splitting by CSSG method (a) original image (b) segmented image

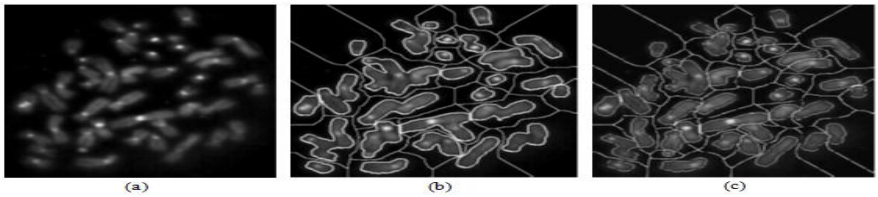

Fig.5.Results of recursive watershed transform technique (a) Original image (b) initial segmentation (c) final segmentation after 2 iterations

Figure 6 shows the segmentation of chromosomes by the 'valley searching' technique imposed by Ji (1994). Valley searching technique attempts to find a valley of gray values corresponding to separation between touching groups of chromosomes. Initially, along the boundary of chromosomes all high concavity points are detected. Next, minimum density path between touching chromosomes is detected by performing a heuristic search. Even though these methods state a high rate for segmenting touching and overlapping chromosomes they use large number of parameters. Also the pale path feature is inconsistent when two chromosomes touches very closely or partially overlapped, as the splitting path is not visible and there is no clear indentation. Figure 6 (a) shows the segmentation of touching chromosomes: (a) shows the cluster of touching chromosomes and its boundary and convex hull. c1, c2 and c3 mark the concavity regions and the potential cut points; (b) shows the pale path and (c) the resolved cluster. Figure 6 (b) shows the segmentation of overlapping chromosomes (a) shows the original overlapped chromosome images and (b) is the decomposed images.

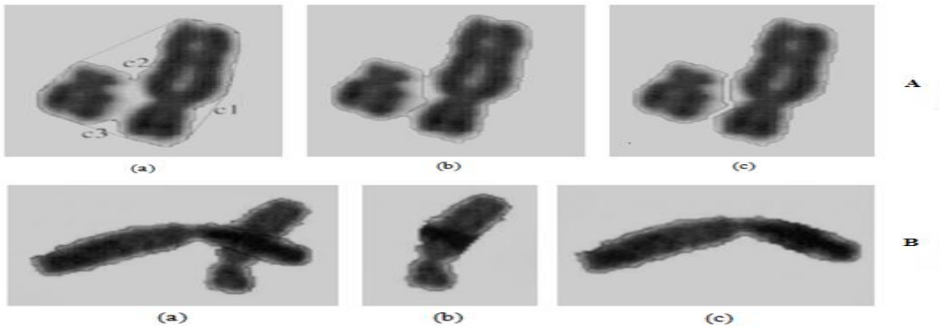

Fig.6.Segmentation results of valley searching technique

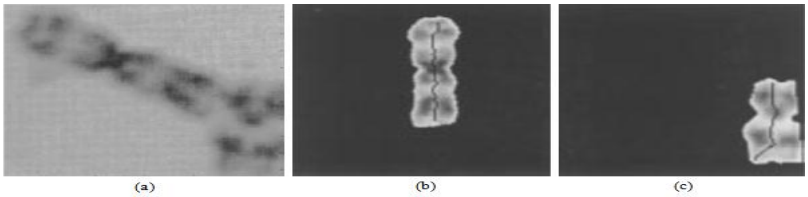

Fig.7.Segmentation results of CPOOS method a) original image b-c) segmented images

Figure 7 shows the segmentation result obtained by Lerner (1998), CPOOS method. CPOOS is a classification driven method for segmenting partially occluded objects that combines a simple segmentation stage and recognition phase. It uses object curvature and intensity image. The object which creates the occlusion is identified by incorporating 'apriori' information. This information is used to set the number of classifier outputs which allows a lower complexity network and shorter training sessions. This method only handles the image with only one cluster and clusters with only two objects. The first image of figure 7 shows the original image of touching chromosomes and the next images are the segmented images. The profiles are measured along the medial axis.

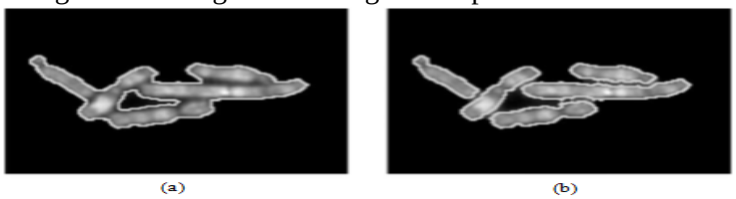

Fig.8.Segmentation result obtained by using geometric evidence and image information (a) initial touching chromosomes (b) final segmentation

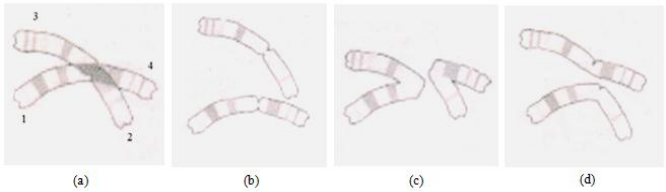

Fig.9.Overlap resolving procedure of Hypothesis and Test strategy

Figure 8 shows the segmentation of touching chromosomes by using 'geometric evidence and image information, technique imposed by Grisan (2009). This method tackles segmentation, overlaps and adjacencies. This algorithm is tested on real images and has reasonable computational complexity. Figure 9 shows how two-chromosome overlap is resolved by 'Hypothesis and Test' strategy imposed by Graham (2002). Figure 9 (a) shows four labelled isolated segments. These can combine to form six hypothesized chromosomes in three different solutions: (b) 1+4; 2+3, (c) 1+3; 2+4 and (d) 1+2; 3+4. The probability of these hypotheses corresponding to genuine chromosomes can be assessed on the basis of shape. The chromosomes in (b) look most likely; those in (d) seem reasonable; those in (c) appear to be the least likely solution. The banding pattern can also be used to assess the hypotheses by matching templates of short lengths of banding pattern to the visible segments. This hypothesis and test method is limited to resolve two chromosome clusters. If same evidence could be used for larger clusters, the complexity of analysis increases with number of chromosomes involved. Again it is evaluated using simulated overlaps and it does not generate objects of unusual appearance.

DISCUSSION

The eight disentangling methods considered in this work are based on five main ideas: pale paths (between touching chromosomes), shape validation, shape concavities, the skeleton and segmentation driven by classification, as already mentioned in section I. Even though every single method has its own weakness and strengths, the ones

belonging to the same category have a tendency to share similar values in performance assessment. Table 1 shows the comparison of various methods for chromosome segmentation found in the literature. Ji's (1989; 1994), method analyses skeleton and boundary based on curvature and convex hull. It forces the image to contain reasonable number of chromosomes (45-47) regardless of their likelihood and of the actual number of chromosomes in image. The splitting phase is based on pale path. It resolves only X-shaped overlaps using geometric evidence provides an interesting comparison with the use of shape evidence. This method is applicable only for unbanded chromosomes and as there are many pale paths in banded chromosomes it is difficult to apply pale path concept.

To overcome this difficulty, Popescu (1999), proposed a technique for pale path cutting that evaluates the resulting candidate chromosomes in hypothesis and test strategy. This study also deals with overlapping chromosomes by assessing hypothesized objects constructed from segments based on their banding pattern. This method analyses the boundary and axis shape that is less heuristic than Ji (1989; 1994). Agam and Dinstein (1997), uses shape evidence for resolving clusters. This approach is applicable to touching or slightly overlapping configurations (T-shaped). It evaluates hypotheses using simple, heuristic shape models. Lerner *et al.* uses [16] banding evidence for resolving clusters. This method assesses the hypothesized chromosomes by classification of banding pattern with an MLP neural network. It combines the choice of correct cluster disentanglement with the classification stage, resulting in classification driven segmentation. Also it is suitable to separate clusters of two chromosomes only. It also has a serious constraint that, classes of chromosomes involved need to be known 'apriori'. After classification of isolated chromosomes, this 'apriori' knowledge is obtained by using simple elimination criterion. But this is not reasonable, because, it needs only one overlap per cell and all the remaining chromosomes are needed to be classified with high accuracy. To overcome the constraint, Popescu (1999), adopts optimization strategy. It achieves lower correct recognition rate for overlapping chromosomes, but for more difficult clusters it applies a much more realistic analysis.

Graham (2002), uses trainable models of shape to resolve difficult cases allowing shape and classification evidence to be combined with probabilistic classification evidence. This approach integrates segmentation and classification to resolve overlaps and adjacencies. It uses split and merge approach in the context of an iterative system to overcome the difficulty of applying pale path technique to banded chromosomes. Ritter and Gao (2001), also combine segmentation and classification, but relies on generation of number of possible chromosome configuration called variants. This method can resolve almost all type of overlapping configurations. But, it has the following disadvantages: a) it assumes that, there is a limited and known set of configurations onto which the chromosomes may be arranged, b) use of heuristic set of rules to identify a cluster from a single chromosome, c) assumes that image contains the whole set of cell chromosomes, d) high number of variants taken into account, which shows great difficulty in resolving clusters with high number of chromosomes involved, e) high number of parameters to set and long computational time.

Table 2 shows the number of chromosomes used by the authors for testing their individual techniques and the accuracy obtained. Authors use more number of chromosomes for testing their algorithm, but they attained low accuracy compared to others. This low accuracy is due to the presence of wrong cuts in the clear and ambiguous phases and missing cuts in the ambiguous phase. Lerner *et al.* [16] uses only few hundreds of chromosomes, but they attained comparatively better performance. Figure 10 is the plot of number of chromosomes used for testing against accuracy.

Table.1.Comparison of various chromosome disentanglement methods found in the literature

Communication	Popescu	Ji	Ji	Grisan	Graham	Ritter & Gao	Karvelis	Lerner
Pale path	*	*	*	*			*	
Concavities	*	*	*			*		
Skeleton			*	*		*		
Shape validation	*			*	*	*		
Segmentation driven by classification					*	*		*
banded	*	*		*	*	*		*
unbanded			*				*	
overlap	*		X-shaped	*	*	all		Slightly (T-shaped)
touching		*		*			*	

Table.2.Number of chromosomes used for testing and accuracy comparison

No. of chromosomes	3979	11279	19719	6683	6177	28148	940	481
accuracy	83	95	91-95	94	86.2	55	92	82.6

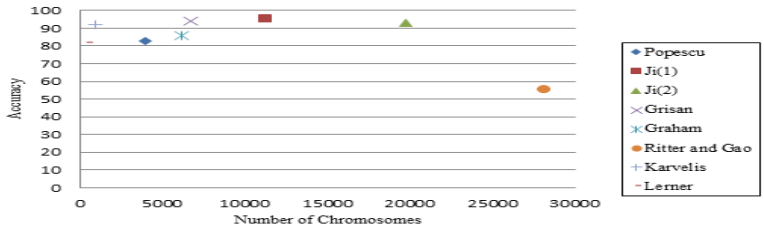

Fig.10.Plot of number of chromosomes against accuracy

4. CONCLUSION

In this paper, we have considered eight disentangling methods for chromosome image segmentation. Each one has its own strong and weak points. Valley searching and recursive watershed algorithms are suitable only for unbanded chromosomes. Most authors resolves only touching chromosomes or slightly overlapping chromosomes. Even if they deal overlapped chromosomes, the method is tested using simulated overlaps. For resolving touching chromosomes, pale path technique is commonly used. Besides all these method, variant analysis technique resolves almost all types of overlap configurations, but gives lower accuracy compared to other methods. This low accuracy is mainly because of wrong cuts and missing cuts. In terms of accuracy, the techniques 'boundary curvature and pale path', 'valley searching' and 'geometric evidence and image information' have almost same result. But geometric evidence and image information technique is considered to be best, as it resolves both touching and overlapping chromosomes.

REFERENCES

Agam G, and Dinstein I, Geometric separation of partially overlapping nonrigid objects applied to automatic chromosome segmentation, IEEE Trans. Pattern Anal. Mach. Intell., 19 (11), 1997, 1212–1222.

Graham C Carters, and Graham J, Disentangling Chromosome overlaps by combining Trainable shape Models with Classification Evidence, IEEE trans. on sig. processing, 50 (8), 2002.

Grisan, Poletti E, and Ruggeri A, Automatic Segmentation and Disentangling of Chromosomess in Q-band Prometaphase Images, IEEE trans. on information technology in Biomedicine, 13 (4), 2009.

Ji L, Fully automatic chromosome segmentation, Cytometry, 17, 1994, 196–208.

Ji L, Intelligent splitting in the chromosome domain, Pattern Recognit., 22 (5), 1989, 519–532.

Karvelis P.S, Segmentation of chromosome images based on a recursive watershed transform, 3rd Eur. Med. Biol. Eng. Conf., EMEBC 2005, Prague, Czech Republic, 2005.

Lerner, Guterman H, and Dinstein I, A classification-driven partially occluded object segmentation (CPOOS) method with application to chromosome analysis, IEEE Trans. Signal Process., 46 (10), 1998, 2841–2847.

Popescu M, Automatic karyop typing of metaphase cells with overlapping chromosomes, Comput. Biol. Med., 29, 1999, 61–82.

Ritter G, and Gao L, Automatic segmentation of metaphase cells based on global context and variant analysis, Pattern Recognit., 41, 2008, 38– 55.

Somasndaram D, Vijay Kumar V.R, Separation of overlapped chromosomes and pairing of similar chromosomes for karyotyping analysis, Measurements, 48, 2014, 274-281.

Yongqiang Zhao, Joint segmentation and pairing of multispectral chromosome images, Pattern Anal. Applic., 16, 2013, 497-506.