

Feature Analysis and Centromere Segmentation of Human Chromosome Images Using an Iterative Fuzzy Algorithm

Parvin Mousavi, Rabab Kreidieh Ward*, *Fellow, IEEE*, Sidney S. Fels, Mohammad Sameti, and Peter M. Lansdorp

Abstract—Classification of homologous chromosomes is essential to advanced studies of cancer genetics. Centromere intensities are believed to be an important differentiating feature between homologs. Therefore, segmentation of centromeres is a major step toward the realization of homolog classification. This paper describes an iterative fuzzy algorithm which successfully segments centromeres from images of human chromosomes prepared using fluorescence *in-situ* hybridization technique. The algorithm is based on assigning a fuzzy membership value to each pixel in the centromere image. An iterative algorithm then updates and minimizes a defined error function. Chromosome 22, a highly heteromorphic chromosome, is used to verify the centromere segmentation method. Homologs of this chromosome are classified based on their segmented centromere intensities as well as their morphological differences. The classification results of these two methods agree completely and are used to validate our developed algorithm.

Index Terms—Centromere segmentation, fuzzy logic, homolog classification, human chromosomes.

I. INTRODUCTION

AT present, advances in cancer research are largely dependent on developments in image processing techniques. This is partly due to the advent of fluorescence *in-situ* hybridization (FISH) technology which makes acquiring microscopy images of human chromosomes with quantitative information feasible.

The nucleus of cells in the human body are made up of 23 pairs of chromosomes. In each pair, one chromosome is inherited from the mother known as the maternal homolog and the other chromosome is inherited from the father known as the paternal homolog. Scientists believe cancer is related to specific

Manuscript received January 10, 2001; revised November 5, 2001. This work was supported in part by the Faculty of Graduate Studies, University of British Columbia and in part by the National Institutes of Health (NIH). *Asterisk indicates corresponding author.*

P. Mousavi was with the Department of Electrical and Computer Engineering, University of British Columbia and the Terry Fox Laboratory for Hematology/Oncology, B.C. Cancer Research Centre, Vancouver, BC 6 1Z4, Canada. She is now with Molecular Mining Inc., Kingston, ON K7L 2Y4, Canada.

*R. K. Ward is with the Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC V6T 1Z4, Canada (e-mail: rababw@ece.ubc.ca).

S. S. Fels is with the Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC V6T 1Z4, Canada.

M. Sameti is with BrainTech Inc., Vancouver, BC V6T 1Z4, Canada.

P. M. Lansdorp is with the Terry Fox Laboratory for Hematology/Oncology, B.C. Cancer Research Centre and Department of Medicine, University of British Columbia, Vancouver, BC V6T 1Z4, Canada.

Publisher Item Identifier S 0018-9294(02)02993-2.

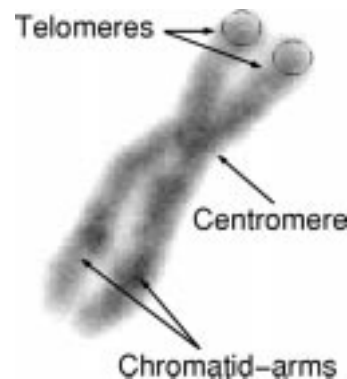


Fig. 1. Structure of a human chromosome.

chromosome abnormalities. To study the characteristics of cancerous cells, it is essential to have the ability to analyze separate homologs [1].

The structure of a metaphase chromosome is shown in Fig. 1. As seen from this figure, a metaphase chromosome consists of two sister chromatids with four arms, known as *P* arms and *Q* arms. The sister chromatids are simply duplicates and identical most of the time. The center of the chromosome is known as the “Centromere” and “Telomeres” form the ends of the *P* and *Q* arms. Telomere lengths are believed to have an important role in cell life-span and development of cancerous cells. Cancer is known as a somatic genetic disease, in which an abnormal form of a gene suddenly appears in some part of the body. The first acquired genetic abnormality provides some growth advantage to the cells. Upon each round of cell division, telomeres shorten until one telomere reaches a critical length. At this point a DNA damage signal is generated and the abnormal cells stop dividing. It seems that in different individuals, different chromosomes become shorter sooner than others, since specific chromosomes vary in telomere lengths [2]. As specific chromosomes have critically short telomeres, they also become unstable and may fuse with other short telomeres. This may trigger loss of genetic loci (the specific place on a chromosome where a gene is located) in chromosomes, known as loss of heterozygosity (LOH). LOH is a major manifestation of the genetic instability that is a characteristic of cancer cells.

To test the role of specific chromosomes in LOH, the ability to analyze telomere lengths in separate homologs would greatly increase the resolving power of the analysis. Moreover, even though the somatic changes that occurred and are not passed on to the next generation, yet various predispositions to cancer

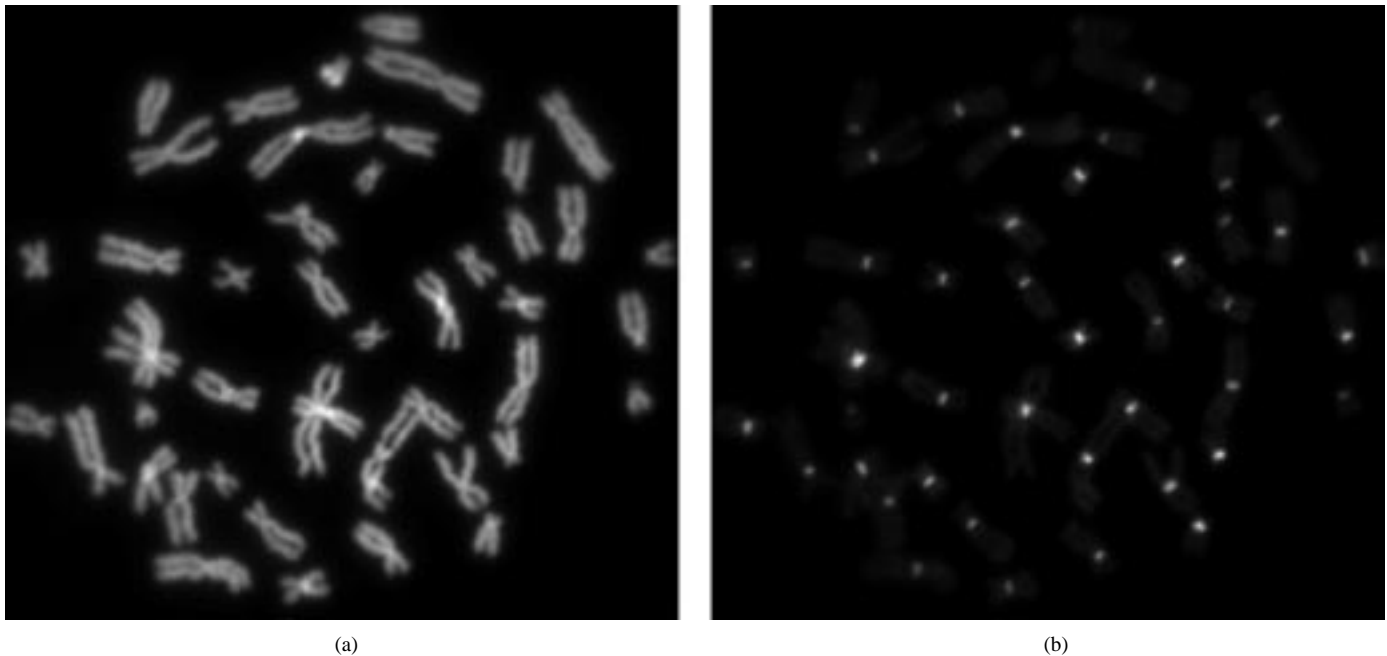


Fig. 2. Snapshots of a metaphase image prepared using (a) DAPI- and (b) FITC-centromere probes.

are inherited as abnormal genes. In addition, in certain cases, the statistics also indicate the possibility of cancer being an inherited genetic disease. In order to get reliable quantitative information on cancer as an inherited genetic disease, it is important to differentiate between maternal and paternal homologs and trace the genes responsible for cancer back into past generations. Until now, all chromosome specific telomere length analysis have been done on pooled homologs.

Despite the vast amount of research conducted on analyzing chromosome images for karyotyping purposes [3]–[5], it has only been a few years since the problem of homolog classification (into maternal and paternal classes) has been addressed [1], [2], [6], [7]. Microscopy images of chromosomes available in the past did not contain enough quantitative information. Recently, however, with the development of novel peptide nucleic acid (PNA) probes, some at the Terry Fox laboratory of the BC Cancer Research Center (BCCRC), chromosome images with high quantitative information are provided [8].

In this paper, a new algorithm is proposed for centromere (central part of chromosome) segmentation. The algorithm is based on fuzzy set theory and gradient descent method. Centromeres of human chromosomes are segmented and distinguished from their background using this algorithm. An intensity feature is then calculated from the segmented centromere areas. Homologs of chromosome 22 are next classified into two classes of maternal and paternal chromosomes based on the differences in their calculated centromere intensity feature. There is no biological method to verify the segmentation results; however, for the specific chromosome 22, paternal and maternal homologs are differentiable due to morphological variations. We use this information to validate our segmentation algorithm. Homologs of chromosome 22 are classified based on the results of our segmentation algorithm as well as their morphological differences. The results of the two classification methods are compared to confirm our segmentation algorithm.

The layout of the paper is as follows. In Section II, we give an overview of the data-base preparation and image acquisition procedures. Section III discusses the centromere segmentation algorithm. In Section IV, we present the results of this algorithm as well as homolog classification for chromosome 22. The conclusion of the paper is presented in Section V.

II. IMAGE ACQUISITION AND DATABASE PREPARATION

A slide of metaphase chromosomes is treated with certain types of PNA and DNA probes. Probes are fluorescent nucleic acid segments that bind to specific sites (substructures) on chromosomes and make these sites observable. Twelve images of the prepared slide are then taken by a fluorescence microscope (Widefield Zeiss Axioplan) and preprocessed prior to use. As probes are fluorescent, only the strings the probes bind to are visualized by the fluorescence microscope [1]. The probes used for preparing our database are 4'-6-Diamidino-2-phenylindole (DAPI) and Fluorescein isothiocyanate-centromere (FITC) probes. DAPI has a blue appearance, highlights the whole chromosome, and is excited at a wavelength of 405 nm. FITC, on the other hand, has a green appearance, makes only the centromeres of the chromosomes visible and is excited at a wavelength of 490 nm (see Fig. 2). In preparing multiprobe images, care is taken to choose probes that have little spectral overlap. In addition, to further minimize the effect of this overlap, multiple filters are used prior to image acquisition. However, there still exists some spectral leakage of the DAPI image into other image planes.

The image acquisition system is divided into three subsystems as follows: 1) A fluorescence microscope which produces the images of the prepared slide; 2) A CCD camera to capture and digitize the image; and 3) A computer system which stores the images, sends instructions to the microscope stage and is used in the preprocessing of the stored images. Preprocessing is

an important stage in preparing our images. Acquired images of a single slide may differ from one another due to a number of factors such as the preparing conditions of the slide, the location of the slide in the field of view of the microscope, the illumination stability and photobleaching. Preprocessing eliminates most of these effects and enables us to have consistent output images. The algorithms used at this stage include background subtraction, wavelength compensation, and photobleaching correction methods [6].

Once images of metaphase chromosomes have been acquired, karyotyping is performed using a semi-automatic procedure. Karyotyping is the process of recognizing and categorizing the 23 different chromosome types from images of metaphase spreads. For karyotyping, we use a software from Metasystems corporation [9]. This software partially automates the karyotyping process and a technician then completes the procedure. The software also generates a table that contains the coordinates of all chromosomes and their karyotype numbers. The data-base is created by cutting out (isolating) every homologous chromosome pair (from the DAPI image) as well as its corresponding centromeres (from the FITC image) and forming separate images for each. We developed an automated algorithm for this purpose. The algorithm uses the table of coordinates of chromosomes to find and cut out each chromosome and its corresponding centromere as a rectangular region. These rectangular regions are saved as images that make up our data-base. The objective is to segment all the centromere images from their background and extract features from the segmented centromeres for homolog classification (specifically homolog classification of chromosome 22) purposes.

III. CENTROMERE SEGMENTATION

Although centromeres in the FITC images have higher intensities than their backgrounds, segmentation is not a straightforward task. The boundaries of centromeres in microscopy images are not well defined and have extremely gradual transitions; therefore, simple thresholding methods are not successful in segmenting them from their backgrounds. In addition, the techniques for preparing and acquiring chromosome images are changing rapidly. This in turn results in centromere image databases that are acquired under different conditions with variable qualities, intensities and contrasts. Human beings seem to be able to resolve objects in imperfect or variable images such as centromere images. To perceive objects in imperfect images, it appears that humans are applying heuristic algorithms to understand such images [10]. In order to systematically characterize the heuristic processes implemented by the human visual system, we suggest fuzzy segmentation. Fuzzy centromere segmentation is a suitable framework for expressing the heuristic process applied to variable and imperfect data such as centromere images. As the centromeres have unclear boundaries, the fuzzy set approach makes the process of assigning a pixel to the centromere or the background region more accurate. In addition, sometimes, there exists a small spectral overlap of DAPI images in the FITC-centromere image. In order to determine whether or not a pixel belongs to the centromere, we need to examine the

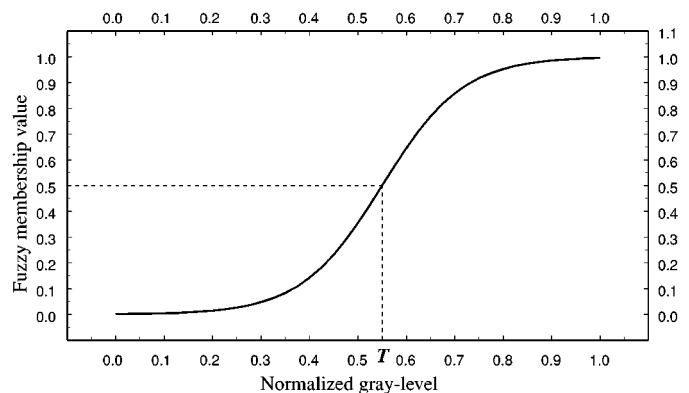


Fig. 3. Fuzzy membership function.

intensities of the surrounding pixels as well as the gray level value of that particular pixel. A neighborhood mask is used for this purpose. Centromere segmentation is performed in three steps described in Sections III-A–C.

A. Image Normalization and Neighborhood Mask

The acquired microscopy images of centromeres are indexed images with a gray-level range of 0–255. Each image is initially converted to an intensity image by determining the minimum and maximum gray levels of the image and normalizing all the pixel values in the image to a number between 0.0 and 1.0.

As mentioned above, for determining the membership of a pixel to the centromere, the surrounding pixels of that particular pixel need to be examined. Therefore, a 3×3 neighborhood mask, centered at the pixel of interest, with distinct weight coefficients is defined as

$$W(i, j) = \exp\left(-\frac{\|(i, j)\|}{\alpha}\right) \quad (1)$$

where $\|(i, j)\|$ is the Euclidean distance between position (i, j) and the center of the mask [11]. The constant α controls the shape of the exponential and is set to two.

B. Fuzzy Membership Values

We define a fuzzy set \mathbf{A} as

$$\mathbf{A} \triangleq \{\text{centromere pixels of an FITC image}\}. \quad (2)$$

A sigmoid fuzzy membership function is applied to the centromere image assigning a membership value to set \mathbf{A} for each pixel of the image (Fig. 3).

The fuzzy membership value of each pixel determines its degree of membership to the centromere (\mathbf{A}) or to the background ($1 - \mathbf{A}$). In Fig. 3, for example, an image pixel with a gray level intensity of 0.9 will be member of the centromere with a probability of more than 90%. On the other hand, a pixel with a gray level intensity of 0.1 is part of the centromere with a probability of less than 5%, that is a background pixel with a probability of 95%.

In the sigmoid membership function, the threshold, T , is always the gray-level value corresponding to the membership value of 0.5. In other words, the location of T relative to the function curve is fixed, i.e., a larger T shifts the membership function to the right while a smaller T shifts the membership

function to the left. For segmentation purposes, usually the threshold, T , is chosen in the valley between the two peaks of the image histogram[12]; however, for our images this is not applicable. The histograms of centromere images do not have two peaks. This is due to the low number of centromere pixels in the image as well as the gray level distribution of the centromere images. In our application, we choose T at a gray level corresponding to the lowest number of pixels, right after the peak of the image histogram (local minimum). In a database of centromere images of chromosomes, this threshold is calculated for each centromere image and the calculated threshold values are averaged to one value T . Therefore, in the database of chromosome 22 images, T is calculated by averaging the threshold from all centromere images in the database. By averaging the values, we determine $T = 0.3$ for this case.

C. Error Calculation and Updating

After assigning a membership value to each pixel, an iterative process updates and finally determines the membership of each pixel to the centromere. The error function is defined as [11]

$$E = \sum_{i,j} O(i,j)(1 - O(i,j)) \quad (3)$$

where $O(i,j)$ is the fuzzy membership value of the image pixel (i,j) . The error, E , is a nonnegative number and is minimized only if the fuzzy membership value, $O(i,j)$, of each image pixel is either one or zero. Inspired by the gradient method and back-propagation algorithm, the updating rule is written as

$$O^+(i,j) = O(i,j) + \eta \left(-\frac{\partial E(i,j)}{\partial O(i,j)} \right) \cdot O(i,j)(1 - O(i,j)) \quad (4)$$

where $E(i,j)$ is the error caused by pixel (i,j) and η is the learning rate. In order to include the role of neighboring pixels in defining a pixel as centromere, using (1), (4) is rewritten as

$$O^+(i,j) = O(i,j) + \eta \left[\sum_{k,l \in N^0} W(k,l) \left(-\frac{\partial E(k,l)}{\partial O(k,l)} \right) \right] \cdot O(i,j)(1 - O(i,j)) \quad (5)$$

where $W(k,l)$ are the weight coefficients of the mask defining a neighborhood N^0 around (i,j) . After substituting (3), the latter becomes

$$O^+(i,j) = O(i,j) + 2\eta \left[\sum_{k,l \in N^0} W(k,l)(O(k,l) - 0.5) \right] \cdot O(i,j)(1 - O(i,j)). \quad (6)$$

Note that pixels with fuzzy membership values close to "0" (background) or "1" (centromere) are highly unlikely to become members of region "1" (centromere) or "0" (background). Therefore, the updating process can be accelerated by applying this information. To do so, (6) is changed to

$$O^+(i,j) = O(i,j) + \Delta O(i,j) \quad (7)$$

where the formula shown at the bottom of the page holds. Equation (7) makes a quick decision for pixels with fuzzy values close to zero or one (because $(O(k,l) - 0.5)$ has a large value), but for a fuzzier pixel ($(O(k,l) - 0.5)$ closer to zero), there is a smaller change in its intensity, causing a delay in segmentation of the pixel. Therefore, for pixels with gray level intensities close to the original threshold (T), the algorithm has high tolerance and precision in segmenting them with respect to T .

IV. RESULTS AND DISCUSSION

As described in Section III, a segmentation algorithm was developed based on fuzzy sets theory and error back-propagation. The algorithm was implemented and used to segment the centromeres from their background for all chromosomes in the prepared data-base. The size of the neighborhood mask $W(i,j)$ is an important variable in the segmentation algorithm. Several experiments were performed with different mask sizes (3×3 , 5×5 and 7×7). As the size of the neighborhood mask increases, the area of the segmented centromere decreases. This is due to the fact that with a larger mask size more bright pixels are needed around a certain pixel for it to be segmented as centromere. Therefore, some edge pixels in the FITC-centromere image will be eliminated during segmentation. The 3×3 neighborhood mask was the best size and was chosen as the algorithm default. Segmentation results, for a pair of homologous chromosome 22, using different neighborhood mask sizes are shown in Fig. 4.

Segmentation was performed successfully on all of the centromere images in the database. Examples of the centromere segmentation for two homologs of chromosome 22 and chromosome 19 is shown in Fig. 5. The original, normalized, and segmented centromeres for the two homologs of chromosome 22 and chromosome 19 are shown in the top and bottom parts of this figure, respectively.

As mentioned before, there is no biological method to validate our segmentation results. In addition, the PNA centromere probes are quite new and manual segmentation of the centromeres by a cytotechnician is not a standard procedure yet. However as a crude verification, we first compared our segmentation results with those of manual segmentation of a cytotechnician on the prepared database. To further verify our results, we used chromosomes 22 as for these chromosomes it is possible to classify their paternal and maternal homologs using differences in their DAPI images. Segmentation results (using our method above) of the FITC images are also used to classify chromosomes 22 into two classes. This is done by calculating an intensity feature, the integrated fluorescence intensity (IFI), over the resultant segmented area of every chromosome and studying the differences in the values of this feature. The classification results of the two methods should agree.

$$\Delta O(i,j) = \begin{cases} 0, & O(i,j) \leq 0.15 \text{ or } \geq 0.9 \\ 2\eta \left[\sum_{k,l \in N^0} W(k,l)(O(k,l) - 0.5) \right], & \text{otherwise} \end{cases}$$

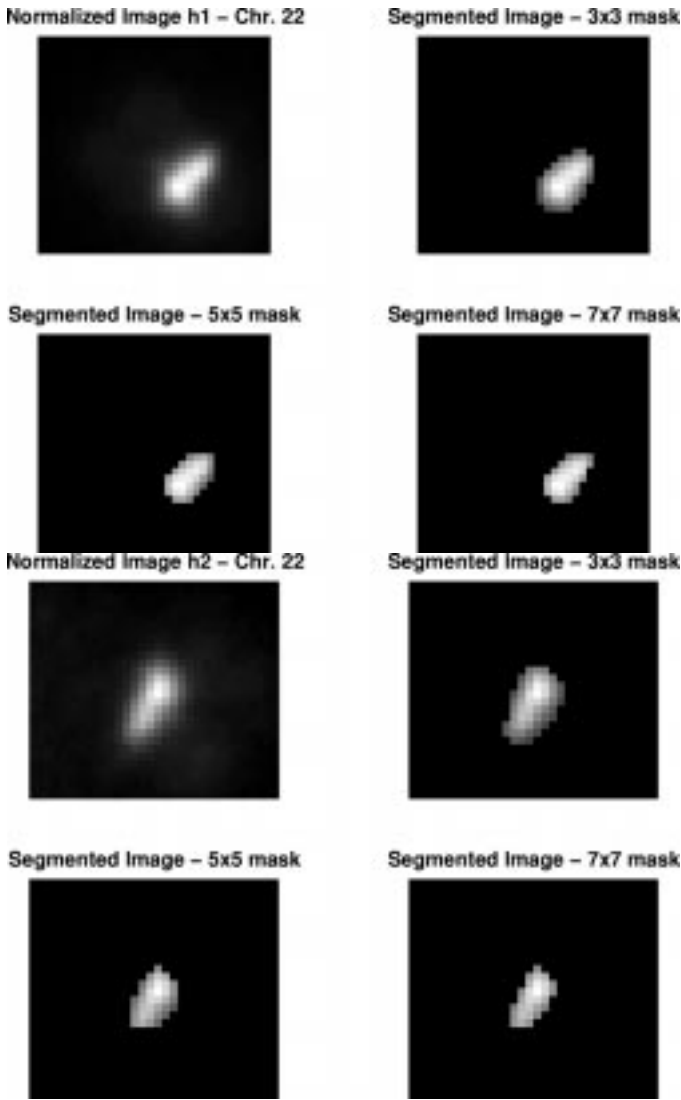


Fig. 4. Segmentation of centromeres for a pair of homologous chromosome 22 using three different neighborhood mask sizes.

A. IFI

The overall intensity seems to be a natural feature to be computed for the segmented centromere. All of the quantitative intensity measurements carried out on the chromosome images are indeed measurements of the IFI values. IFI is a measure of the total amount of fluorescence emitted from an object $O(x, y, z)$ [6], in other words

$$\text{IFI} = \int_x \int_y \int_z O(x, y, z)(dx)(dy)(dz). \quad (8)$$

In [1], we show that the IFI of an object is proportional to the IFI of the observed three-dimensional (3-D) image, $i(x, y, z)$ of that object, that is

$$\text{IFI} = K \int_x \int_y \int_z i(x, y, z)(dx)(dy)(dz). \quad (9)$$

The 3-D image of an object is reconstructed from a series of two-dimensional (2-D) images taken from that object, at dif-

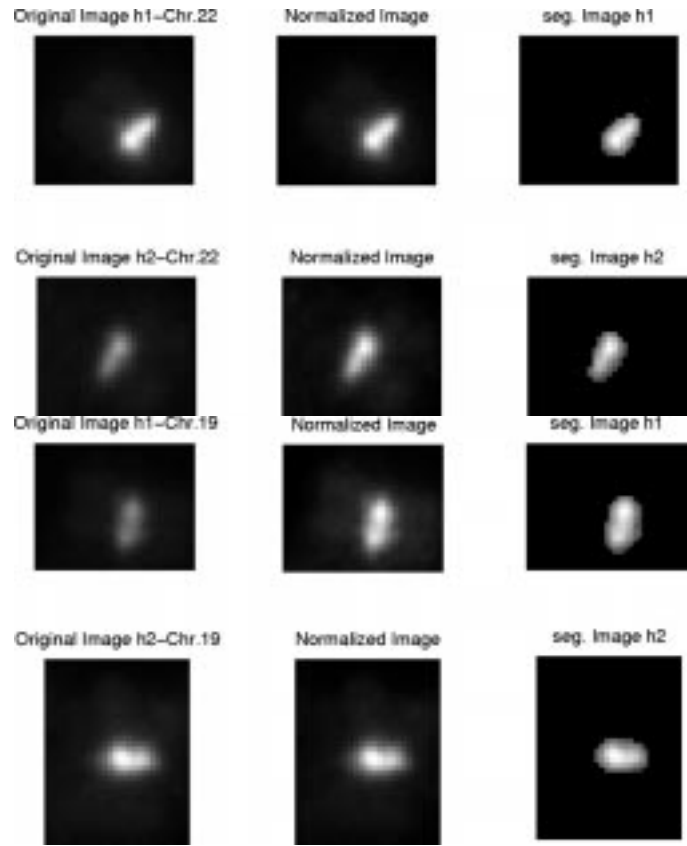


Fig. 5. Centromere segmentation for homologs of chromosome 22 (top) and chromosome 19 (bottom) in a metaphase spread.

TABLE I
THE IFI (OVER THE SEGMENTED AREA) VALUES FOR SEGMENTED FITC-CENTROMERES OF HOMOLOGOUS CHROMOSOMES 22 FROM THE DATABASE (NOT-NORMALIZED VALUES)

Chromosome 22	Image1	Image2	Image3	Image4
with P-arm	11881	10712	7157	6421
without P-arm	7356	7525	4687	5166
Chromosome 22	Image5	Image6	Image7	Image8
with P-arm	7348	5619	7656	8513
without P-arm	5403	4181	5134	4913
Chromosome 22	Image9	Image10	Image11	Image12
with P-arm	8433	7991	9818	7337
without P-arm	5029	4753	5206	4827

ferent z positions on the focal axis. In addition, the *total intensity* of the captured chromosome images at different z -focal levels, $\int_{x,y} i(x, y, z_i)$, are all equal [1], [6]. Hence, the total light intensity over the two planes placed at z_1 and z_2 are equal, that is

$$\int_x \int_y i(x, y, z_1)(dx)(dy) = \int_x \int_y i(x, y, z_2)(dx)(dy). \quad (10)$$

From (9) and (10) we, therefore, conclude that IFI is proportional to the integral of intensities of the 2-D observed image, or

$$\text{IFI} \propto \int_x \int_y i(x, y, z_2)(dx)(dy). \quad (11)$$

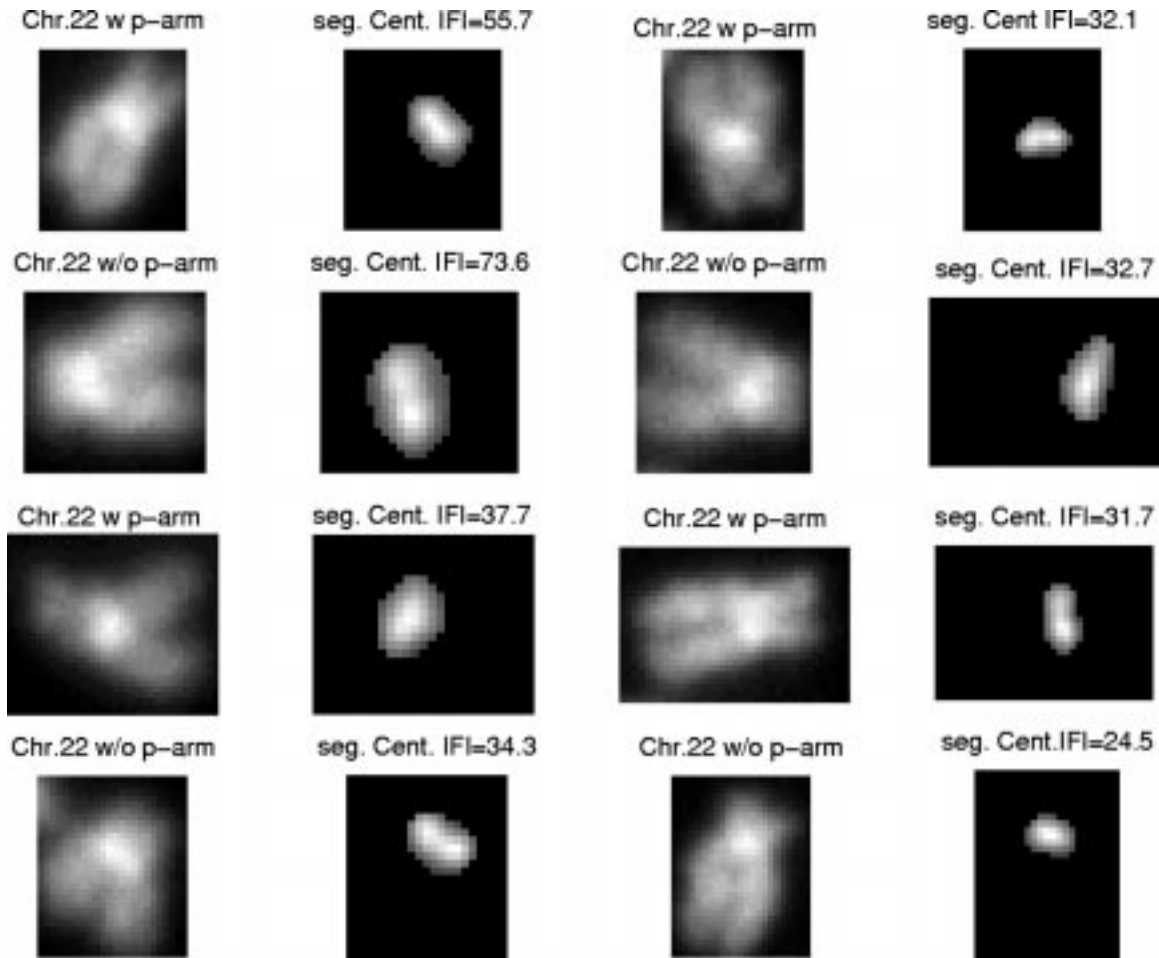


Fig. 6. DAPI and segmented centromeres (from FITC-centromere images) for four pairs of homologous Chromosome 22.

B. Homolog Classification of Chromosome 22

Once the centromere regions in the FITC images are segmented for all chromosomes 22, the total IFI is calculated for each centromere. The IFI values are computed for the two homologs of chromosome 22, for twelve metaphase images of a patient (in the prepared database), using the above method and the results are shown in Table I. Homologs of chromosome 22 are classified into two groups based on the differences in their IFI values.

On the other hand, chromosome 22 is highly heteromorphic. The two homologs of chromosome 22 have apparent differences in their upper arms (P arms) in DAPI images (see Fig. 6). One homolog has short P arms (top left in each section of Fig. 6) while the other homolog has no P arms (bottom left in each section of Fig. 6). Homologs of chromosome 22 are also classified using this heteromorphic difference.

In Fig. 6, examples of four pairs of chromosome 22 homologs with and without P arms as well as their segmented centromeres and total IFI values are shown. As seen from this figure and Table I, homologs with P arms have higher IFI values than homologs without P arms. Homolog classification is performed successfully using both the difference in the P arms and the difference in centromere IFI values. The results of the two methods comply perfectly and the accuracy of the classification is 100%.

Chromosomes 22 are classified into two classes representing maternal and paternal homologs.

A challenge in development of homolog classification algorithms is to overcome the noise introduced during slide preparation and image acquisition stages. Despite the precautions taken in the fore-mentioned stages, our segmentation method should perform well in the presence of acceptable noise. In addition, acquiring large databases of chromosome images is a very lengthy process and can be expensive. In order to provide a larger database but still be able to cross validate our results with ground truth, we added different amounts of random noise to the original centromere images to create new images. The database preparation procedure (including slide preparation, and image acquisition) is a dynamic process, can be variable and includes a great deal of human interaction. Therefore, generating simulated centromere data by modeling the inverse dynamics of the procedure is very difficult if not impossible. In Section IV-C, we model the total noise with a Gaussian distribution function [13] and generate 1200 new centromere images. We then study the performance of the segmentation algorithm on the noisy images. Finally, we cross validate the segmentation results by classifying chromosomes 22 into two parental homolog classes using 1) the centromere intensity differences calculated following segmentation and 2) the differences in chromosomes 22 P arms. Based on this

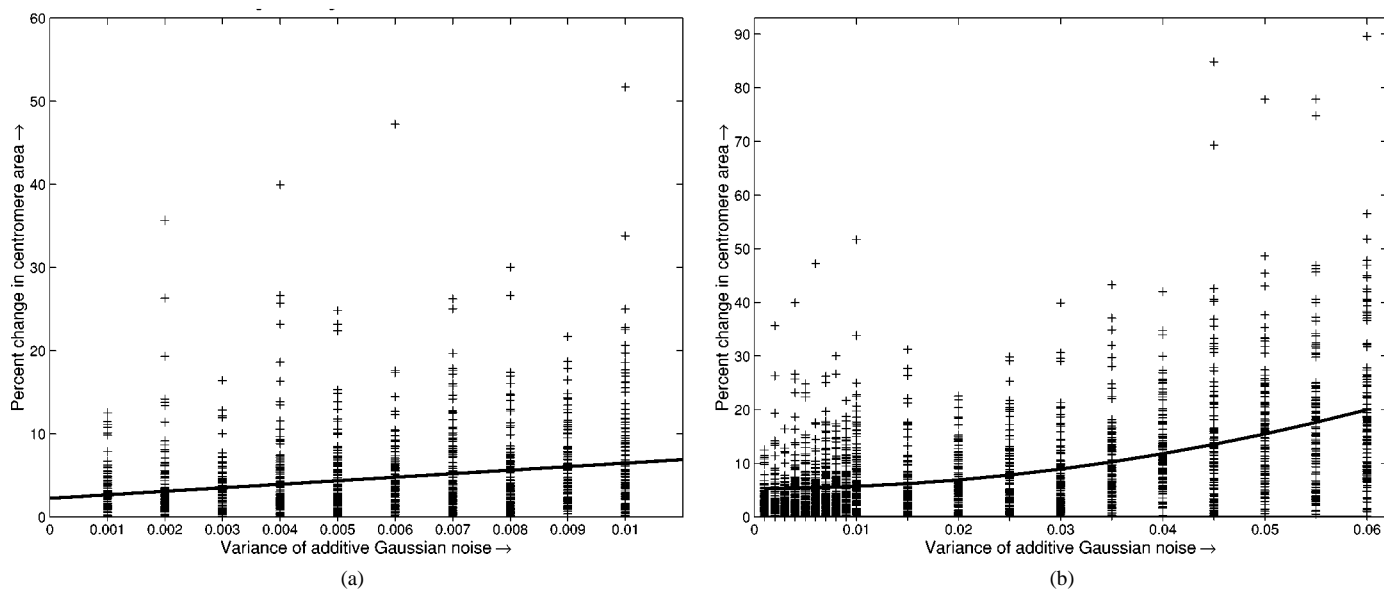


Fig. 7. Scatter-plots of relative changes in segmented centromere areas versus additive Gaussian noise.

cross validation, we can also measure the tolerance of the segmentation and classification algorithms to noise.

C. Simulated Database and Noise Tolerance

Gaussian noise is the simplest form of additive random noise whose probability distribution function is defined as

$$P(x) = \frac{1}{\sqrt{2\pi}\sigma_x} \exp\left[-\frac{(x - \mu_x)^2}{2\sigma_x^2}\right] \quad (12)$$

where μ_X and σ_X^2 are the mean and variance of the Gaussian noise, respectively [13], [14]. In order to generate new centromere images as well as to test the tolerance of our segmentation algorithm to noise, we started with adding Gaussian noise with zero mean and 0.001 variance to our original centromere images. The variance of noise was gradually increased first in smaller steps (0.001) and later in larger steps (0.005 and 0.05). Centromeres were segmented from the noisy images using our segmentation algorithm and the relative changes in the total intensities of the segmented areas were measured. Chromosomes 22 were classified into two classes of maternal and paternal homologs based on the intensity differences of their segmented centromeres. These results were cross validated by classifying homologs of chromosome 22 based on their *P* arm differences as well. In Fig. 7(a) and (b), the scatter-plots for relative changes in segmented centromere area versus additive Gaussian noise are shown for new centromere images. In Fig. 7(a), the variance of noise is between, 0.0 and 0.01. We fitted the best line to the scatter plot in Fig. 7(a). As seen from this figure, the average amount of change in centromere intensity is quite low. This change is about 2.5% for Gaussian noise with 0.001 variance and 6% for Gaussian noise with 0.01 variance. This amount is very low compared with typical centromere differences in homologous chromosomes which is approximately 30% (see Table I). Homologs of most of the newly generated chromosomes 22 (variances up to 0.01) are successfully classified into parental classes using the results of the developed segmentation algorithm and the classification is cross validated.

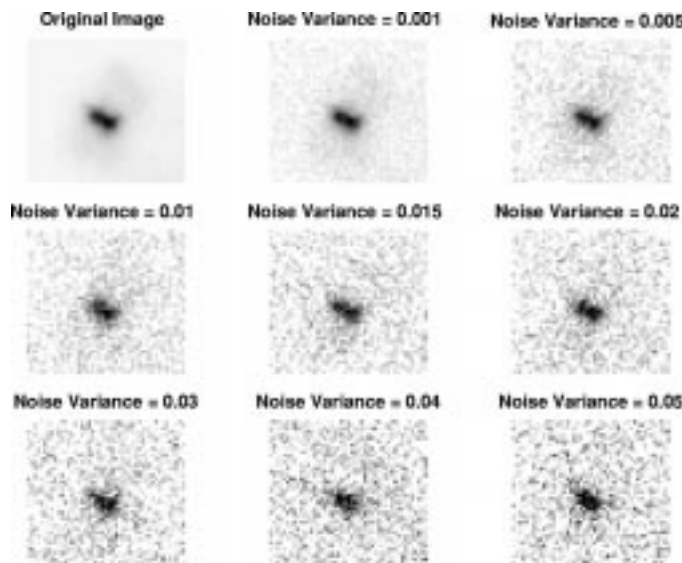


Fig. 8. An example of the original and noisy centromere images for additive Gaussian noise with different variances.

Increasing the amount of the Gaussian noise in the image, the relative change in the centromere intensity increases as well. The scatter plot and polynomial approximation to this plot is shown in Fig. 7(b) for noise variances up to 0.06. Increasing the noise variance up to 0.6, the scatter-plot can eventually be approximated by an exponential function.

An example of a centromere image and the noisy images resulting from additive Gaussian noise with different variances are shown in Fig. 8. As seen in this image, noise with a variance of more than 0.01, distorts the centromere image extensively. The distortion is obvious to the naked eye and the shape of the centromere changes. These images, if produced during the image acquisition procedure, can be easily distinguished and discarded by the technician. Therefore, although high relative-changes occur in segmented centromere areas with noise variance >0.01 , our segmentation algorithm will not be dealing

with these images at all. On the other hand, the segmentation is successful in the presence of low amounts of noise and the changes in the segmented area are acceptable. In addition, the developed algorithm is further evaluated by segmenting the new centromere images and classifying parental homologs of chromosome 22.

V. CONCLUSION

In this paper, an iterative fuzzy segmentation algorithm was implemented and applied to FITC-centromere images of chromosomes. Using this algorithm, the centromeres were successfully segmented from their backgrounds. To verify the segmentation results, the total IFI of the segmented regions were calculated for chromosome 22 and the homologs of this chromosome were classified into two groups of maternal and paternal classes. Since chromosomes 22 show morphological differences between homologs in the DAPI images, we used this criteria to classify chromosome 22 into parental homolog classes as well. The classification results using these two criteria were compared and the two methods complied completely. This proves that our segmentation algorithm was successful in segmenting the centromeres from their backgrounds. The results of this paper constitute a major contribution to multifeature analysis of chromosome images for homolog classification purposes.

Cancer is known as a somatic genetic disease characterized by specific mutations and genetic instability. Studies have shown various predispositions to cancer are inherited as abnormal genes. Statistics also suggest occurrence of some familial cancers. Following the approaches described, the results of this paper will allow classification of homologous chromosomes and will help track the genes responsible for cancer thus lead to a better understanding of the genetics of cancer.

REFERENCES

- [1] P. Mousavi, R. K. Ward, and P. M. Lansdorp, "Feature analysis and classification of chromosome 16 homologs using fluorescence microscopy image," *IEEE Can. J. Elect. Comput. Eng.*, vol. 23, no. 4, pp. 95–98, 1999.
- [2] U. M. Martens, J. M. Zijlmans, S. S. S. Poon, V. Drogoulski, J. Yui, E. Chavez, R. K. Ward, and P. M. Lansdorp, "Short telomere on the P-arm of human chromosome 17," *Nature Genetics*, vol. 18, pp. 76–80, 1998.
- [3] G. H. Granlund, "Identification of human chromosome by using integrated density profile," *IEEE Trans. Biomed. Eng.*, vol. BME-23, pp. 182–192, 1976.
- [4] J. Piper and E. Graham, "On fully automatic feature measurement for banded chromosome classification," *Cytometry Mag.*, vol. 10, no. 3, pp. 242–255, 1989.
- [5] J. Graham, P. Errington, and A. Jennings, "A neural network chromosome classifier," *J. Radiation Res.*, vol. 33, pp. 250–257, 1992.
- [6] S. S. S. Poon, U. M. Martens, R. K. Ward, and P. M. Lansdorp, "Telomere length measurements using digital fluorescence microscopy," *Cytometry*, vol. 36, no. 4, pp. 267–278, 1999.
- [7] P. M. Lansdorp, N. P. Werwoerd, F. M. Van de Rijke, V. D. Little, R. W. Dirks, A. K. Raap, and H. J. Tanke, "Heterogeneity in telomere length of human chromosomes," *Human Mol. Genetics*, vol. 5, no. 5, pp. 685–691, 1996.
- [8] P. M. Lansdorp, V. Dragowska, N. Rufer, T. Brummendorf, S. S. S. Poon, P. Mousavi, T. Duncan, and U. Martens, "Applications of peptide nucleic acid probes in cytometry," presented at the SAC XIX Int. Congr., Colorado Springs, CO, 1998.
- [9] [Online]. Available: <http://www.metasystems.de/>.

- [10] T. Law, K. Yamada, D. Shibata, T. Nakamura, L. He, and H. Itoh, "Edge extraction using fuzzy reasoning," in *Soft Computing for Image Processing*, S. K. Pal, Ed. New York: Physica-Verlag, 2000.
- [11] M. Sameti and R. K. Ward, "A fuzzy segmentation algorithm for mammogram partitioning," in *Digital Mammography'96*, K. Doi, M. L. Giger, R. M. Nishikawa, and R. A. Schmidt, Eds, New York: Elsevier Science B.V., 1996.
- [12] M. J. T. Smith and A. Docef, *A Study Guide for Digital Image Processing*. Carmel, IN: Scientific, 1997.
- [13] E. Kreyszig, *Advanced Engineering Mathematics*, 8th ed. New York: Wiley, 1998.
- [14] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. New York: McGraw Hill, 1991.



Parvin Mousavi received the M.Sc. and D.I.C. degrees in engineering and physical sciences in medicine from Imperial College of Science, Technology and Medicine, London, U.K. and the Ph.D. degree in electrical and computer engineering from the University of British Columbia, Vancouver, BC, Canada, in 2001.

She is currently a computational scientist with Molecular Mining Inc., Kingston, ON, Canada. Her main areas of interest are image processing, medical imaging, pattern recognition, and data mining for predictive toxicology and *in silico* drug screening.



Rabab Kreidieh Ward (S'71–M'72–SM'85–F'99) received the B.Eng. degree from the University of Cairo, Cairo, Egypt, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of California, Berkeley, in 1969 and 1972, respectively.

She is a Professor in the Electrical and Computer Engineering Department and the Director of the Institute for Computing, Information and Cognitive Systems (ICICS), University of British Columbia, Vancouver, BC, Canada. Her expertise lies in digital signal processing and applications to cable TV, HDTV, video compression, and medical images, including mammography, microscopy, and cell images. She holds six patents and has published approximately 200 papers and chapters in scientific books. Many of her research ideas have been transferred to industry, including her invention of a noninvasive testing system for cable TV, a nondestructive system to measure the size of fish for the aquacultural industry, and a compression scheme for photographic images. Her research on medical images has significantly advanced the process of early detection of breast cancer and the measurement of structures in biological cells.

Dr. Ward is a recipient of a UBC Killam Research Prize and is a fellow of the EIC, the Canadian Academy of Engineers, and the Royal Society of Canada. She was the General Chairperson of the 2000 IEEE International Conference on Image Processing



Sidney S. Fels received the B.A.Sc. in electrical engineering from the University of Waterloo, Waterloo, ON, Canada, in 1988 and the Ph.D. and M.Sc. in computer science from the University of Toronto, Toronto, ON, Canada, in 1994 and 1990, respectively.

He was a Visiting Researcher at ATR Media Integration & Communications Research Laboratories, Kyoto, Japan from 1996 to 1997. He has been an Assistant Professor in the Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC, Canada, since 1998. His areas of interest are human-computer interaction, neural networks, intelligent agents, and interactive arts. He currently heads the Human Communications Technology (HCT) Laboratory and is the Acting Director of the Media and Graphics Interdisciplinary Centre (MAGIC) at the University of British Columbia. Some of his research projects include Glove-TalkII, Glove-Talk, Iamascope, InvenTcl, French Surfaces, Sound Sculpting, and the context-aware mobile assistant project (CMAP).



Mohammad Sameti received the B.Sc. degree in electrical engineering from Sharif University of Technology, Tehran, Iran, and the M.Sc. degree in systems design engineering from University of Waterloo, Waterloo, Canada, and Ph.D. degree in electrical and computer engineering from University of British Columbia, Vancouver, Canada in 1998.

He is currently a Senior Vision Engineer at Braintech Inc., Vancouver. His main research interests are biomedical imaging (digital mammography) and vision systems for robot guidance.



Peter M. Lansdorp received the M.D. degree from the Erasmus University in Rotterdam, the Netherlands, in 1976. He received the Ph.D. degree in immunology and experimental hematology from the University of Amsterdam, Amsterdam, The Netherlands, in 1985.

He is a cell biologist studying the role of chromosome ends (called telomeres) in cells in relation to aging and cancer. He has been at the Terry Fox Laboratory at the BC Cancer Agency, Vancouver, Canada, since 1985 and is a Professor in the Department of Medicine at the University of British Columbia, Vancouver. He has published extensively on the properties of purified blood forming stem cells and has developed techniques to measure the length of telomeric DNA repeat sequences. These techniques are now increasingly used worldwide in one of the hottest areas of biology research: “the role of telomeres in aging and cancer.” He has published over 200 papers, reviews and book chapters.

Dr. Lansdorp is on the board of the National Cancer Institute of Canada and speaks regularly at international scientific meetings.