ISSN 1682-296X (Print)
ISSN 1682-2978 (Online)

Bio Technology

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Biotechnology 5 (1): 111-117, 2006
ISSN 1682-296X
© 2006 Asian Network for Scientific Information

Detection of Specific Human Chromosomal Abnormalities Using Discrete Cosine Transform Based Gradient Vector Flow Active Contours

[1]A. Prabhu Britto and [2]G. Ravindran
[1]Center for Medical Electronics, Department of Electronics and Communication Engineering
[2]Faculty of Information and Communication Engineering, Anna University, Chennai 600 025, India

Abstract: In this study, Discrete Cosine Transform (DCT) based Gradient Vector Flow (GVF) active contours are studied and found to have good strength for using as a suitable tool for the detection of specific human chromosomal abnormalities. The DCT based GVF active contours have been investigated by the authors and already established as an efficient boundary mapping tool for human chromosome spread images and also found to be a suitable classification tool for human chromosome images. Further studies on the DCT based GVF active contours show that specific human chromosomal abnormalities can be detected using these active contours. The main strength of this study is that DCT based GVF active contours are being used for the first time for detection of chromosomal abnormalities and they have been found to have good strength for application as a detection tool for chromosomal abnormalities.

Key words: Chromosome, abnormality, detection, gradient vector flow, active contours

INTRODUCTION

An Active Contour (Kass *et al*., 1987) is an energy minimizing spline hat can detect specified features within an image. They have found much use in the segmentation or boundary mapping of objects in images. There are various formulations for Active Contours (McInerney and Terzopoulis, 1996), among which Gradient Vector Flow Active Contours (Prince and Xu, 1996; Xu and Prince, 1998;2000) are found to be much suitable (Britto and Ravindran, 2005a) for boundary mapping chromosome spread images. Among a few variants of gradient vector flow active contours that were tested (Britto and Ravindran, 2005b), discrete cosine transform based gradient vector flow active contours were established as a standardized, efficient tool for boundary mapping of chromosome spread images (Britto and Ravindran, 2005c; 2005d; 2006a). Studies showed that the Discrete Cosine Transform (DCT) based Gradient Vector Flow (GVF) active contours have inherent strength for chromosome image classification (Britto and Ravindran, 2006b). Further studies have yielded valuable insight into the potential of DCT based GVF active contours for detection of specific human chromosomal abnormalities. This study is a continuation of the major research undertaken in cytogenetic studies by using the GVF technique as base mathematical formulation for investigating chromosomal images. The main significance of this study is that GVF active contours is being found to be a suitable tool for detection of chromosomal abnormalities. This study yields better insight into the strength of DCT based GVF active contours as a suitable tool for detection of specific chromosomal abnormalities. This study should be studied in conjunction with prior earlier research (Britto and Ravindran, 2005a; 2005b; 2005c; 2005d; 2006a; 2006b) and other training and classification pattern recognition techniques, for better understanding.

CHROMOSOME IMAGES

A chromosome is an extended DNA molecule containing multiple genes and associated proteins. During mitosis and meiosis, condensed chromosomes form structures that are visible with high-powered light microscopes (http://www.colorado.edu/ MCDB/ MCDB2150Fall/notes00/L0001.html). The chromosome spread images are obtained using the following general procedure. About 5 mL of blood is removed from the patient. If a fetus is being karyotyped, amniotic fluid is removed from the amniotic sac which surrounds the fetus during development. This is done with the aid of a large syringe and ultrasound picturing. There are cells which have come off the fetus in this fluid. The white blood cells are removed from the blood or the living cells are removed from the amniotic fluid. These cells are then cultured in a medium in which they undergo mitosis. Mitosis is

Corresponding Author: A. Prabhu Britto, Center for Medical Electronics,
Department of Electronics and Communication Engineering, Anna University, Chennai 600025, India

stopped at metaphase using chemicals. The cells are then placed onto a slide and spread out. They are viewed under a microscope which is specially adapted with a camera to take a picture of the chromosomes from one of the cells (http://home.earthlin.net/~heinabilene/karyotypes/karyoty.htm).

CHROMOSOME STRUCTURE AND CLASSIFICATION

The human chromosome consists of 22 pairs of autosomes and one pair of sex chromosomes (X and Y chromosome). Idiogram of the human chromosome is shown in Fig. 1.

Each pair of the autosomes and each of the sex chromosomes vary from each other and the absolute length of any chromosome varies depending on the stage of mitosis in which it was fixed. However, the relative position of the centromere is constant, which means that the ratio of the lengths of the two arms is constant for each chromosome. This ratio is an important parameter for chromosome classification (http://arbl.cvmbs.colostate.edu/hbooks/genetics/medgen/chromo/chromosomes.html).

Chromosomes are arranged into seven groups based on size and centromere location (http://www.colorado.edu/MCDB/MCDB2150Fall/notes00/L0001.html). The sex chromosomes X and Y are not grouped and are indicated separately. The following are the groups of chromosomes based on size and centromere location and the sex chromosomes are also given below:

Group A: Chromosomes 1-3 are largest with median centromere.

Group B: Chromosomes 4-5 are large with submedian centromere

Group C: Chromosomes 6-12 are medium sized with submedian centromere

Group D: Chromosomes 13-15 are medium sized with acrocentric centromere

Group E: Chromosomes 16-18 are short with median or submedian centromere

Group F: Chromosomes 19-20 are short with median centromere

Group G: Chromosomes 21-22 are very short with acrocentric centromere.

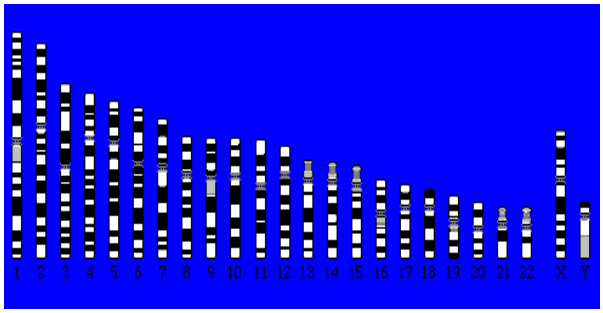

Fig. 1: Idiogram of human chromosomes (Courtesy: http://home.earthlink.net/~heinabilene/karyotypes/karyoty.htm)

Chromosome X: This is similar to group C.

Chromosome Y: This is similar to group G.

The centromeres can be found in the middle of the chromosome (median), near one end (acrocentric), or in between these first two (submedian). A normal karyotype of a male and female are shown in Fig. 2 and 3.

CHROMOSOMAL ABNORMALITIES

The chromosomes in the chromosome spread image are rearranged into homologous pairs, termed as Karyotyping. Such chromosomes are alike with regard to size and position of the centromere (http://anthro.palomar.edu/abnormal/default.htm). A normal male and a normal female karyotype (www.biology.inpui. edu\biocourses\N100\2K2humancsomaldisorders.html) are shown in Fig. 2 and 3.

This rearrangement of chromosomes into homologous pairs allows investigation of chromosomal abnormalities.

The devastating effect of such abnormalities often results in miscarriage during the pregnancy. If the baby survives, it is usually left with a *syndrome*, which is a number of conditions that occur together and characterize a particular disease or condition.

A few syndromes (of interest) are specified below.

An alteration in chromosome number, called *aneuploidy*, is one of the chromosomal abnormalities.

A *trisomic* cell has one extra chromosome, while a *monosomic* cell has one missing chromosome.

Down syndrome (Trisomy 21): Down syndrome arises as a result of trisomy in chromosome 21. A male karyotype with Down syndrome is shown in Fig. 4.

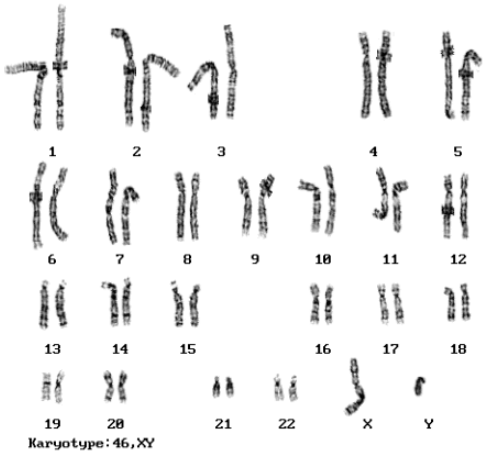

Fig. 2: Karyotype of normal male

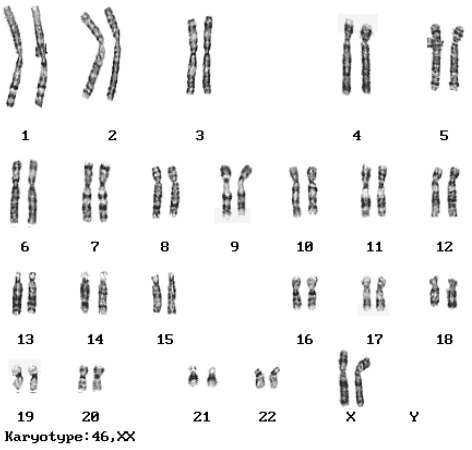

Fig. 3: Karyotype of normal female

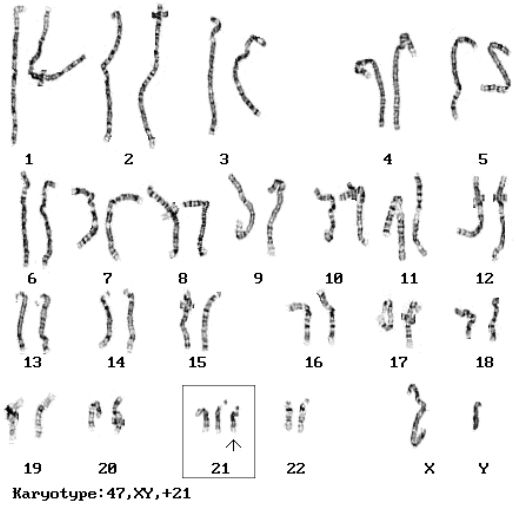

Fig. 4: Male Karyotype with Down Syndrome (Trisomy 21)

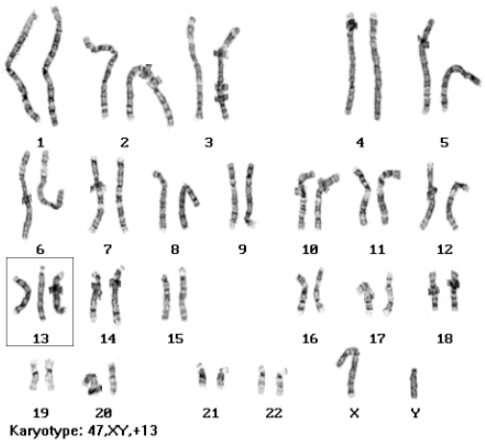

Fig. 5: Male Karyotype with Patau syndrome (Trisomy 13)

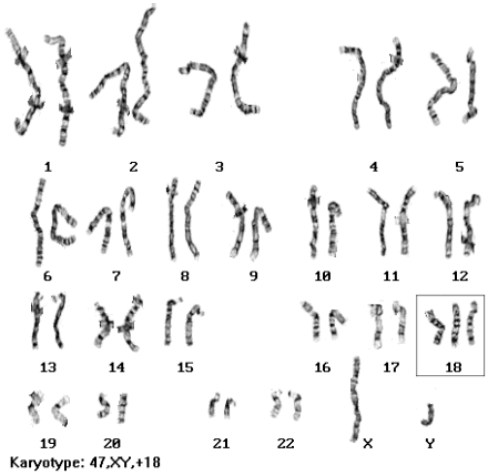

Fig. 6: Male Karyotype with Edward's syndrome (Trisomy 18)

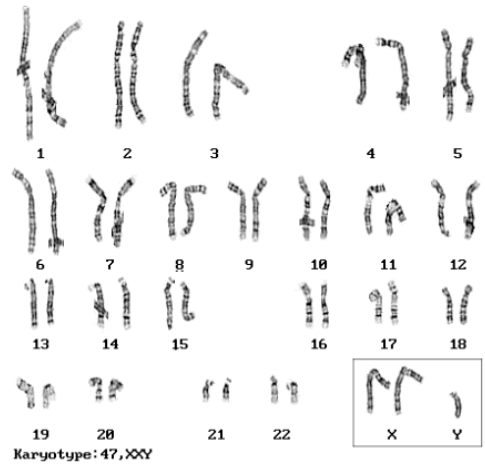

Fig. 7: Karyotype with Klinefelter syndrome (47, XXY)

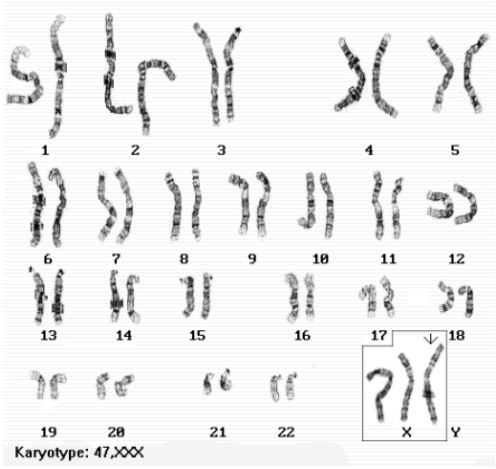

Fig. 8: Karyotype with Trisomy X syndrome

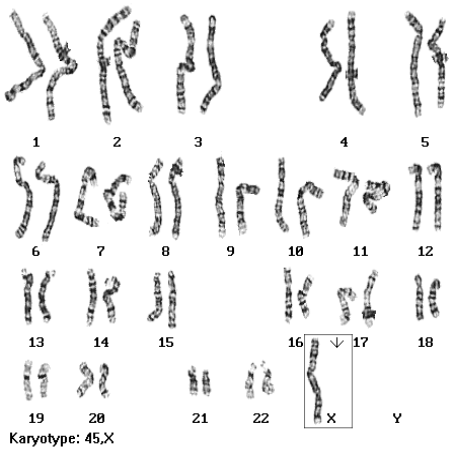

Fig. 9: Karyotype with Monosomy X syndrome
Figure 2-9: Courtesy http://www.biology. iupui. edu/biocourses/N100/2k2humancsomaldisordcr s.htlm

Patau syndrome (Trisomy 13): Patau syndrome arises as a result of trisomy in chromosome 13. A male karyotype with Patau syndrome is shown in Fig. 5.

Edward's syndrome (Trisomy 18): Edward's syndrome arises as a result of trisomy in chromosome 18. A male karyotype with Edward's syndrome is shown in Fig. 6.

Klinefelter syndrome (47, XXY male): Klinefelter syndrome arises as a result of an extra X chromosome in the male karyotype. A karyotype with Klinefelter syndrome is shown in Fig. 7.

Trisomy X (47, XXX female): Trisomy X syndrome arises as a result of an extra X chromosome in the female karyotype. A karyotype with Trisomy X syndrome is shown in Fig. 8.

Monosomy X(Turner's syndrome): Turner's syndrome arises as a result of absence of one X chromosome in the female karyotype. A karyotype with Turner's syndrome is shown in Fig. 9.

GRADIENT VECTOR FLOW (GVF) ACTIVE CONTOURS

Gradient Vector Flow (GVF) Active Contours use Gradient Vector Flow fields obtained by solving a vector diffusion equation that diffuses the gradient vectors of a gray-level edge map computed from the image. The GVF active contour model cannot be written as the negative gradient of a potential function. Hence it is directly specified from a dynamic force equation, instead of the standard energy minimization network.

The external forces arising out of GVF fields are non-conservative forces as they cannot be written as gradients of scalar potential functions. The usage of non-conservative forces as external forces show improved performance of Gradient Vector Flow field Active Contours compared to traditional energy minimizing active contours (Xu and Prince, 1998;2000).

The GVF field points towards the object boundary when very near to the boundary, but varies smoothly over homogeneous image regions extending to the image border. Hence the GVF field can capture an active contour from long range from either side of the object boundary and can force it into the object boundary.

The GVF active contour model thus has a large capture range and is insensitive to the initialization of the contour. Hence the contour initialization is flexible.

The gradient vectors are normal to the boundary surface but by combining Laplacian and Gradient the result is not the normal vectors to the boundary surface. As a result of this, the GVF field yields vectors that point into boundary concavities so that the active contour is driven through the concavities.

Information regarding whether the initial contour should expand or contract need not be given to the GVF active contour model. The GVF is very useful when there are boundary gaps, because it preserves the perceptual edge property of active contours (kass *et al*., 1987; XU and Prince, 1998).

The GVF field is defined as the equilibrium solution to the following vector diffusion equation (Xu and Prince, 2000):

$$u = g(|\nabla f|)\,\nabla^2 u - h(|\nabla f|)\,(u - \nabla f) \quad (1a)$$

$$u(x, 0) = \nabla f(x) \quad (1b)$$

Where, u_t denotes the partial derivative of u (x,t) with respect to t, ∇^2 is the Laplacian operator (applied to each spatial component of u separately) and f is an edge map that has a higher value at the desired object boundary.

The functions in g and h control the amount of diffusion in GVF. In Eq. 1, $g(|\nabla f|)\nabla^2 u$ produces a smoothly varying vector field and hence called as the smoothing term, while $h(|\nabla f|)(u - \nabla f)$ encourages the vector field u to be close to ∇f computed from the image data and hence called as the data term. The weighting functions g (·) and h (·) apply to the smoothing and data terms, respectively and they are chosen (Xu and Prince, 1998) as $g(|\nabla f|) = \mu$ and $h(|\nabla f|). = |\nabla f|^2$. G (·) is constant here and smoothing occurs everywhere, while h (·) grows larger near strong edges and dominates at boundaries.

Hence, the Gradient Vector Flow field is defined as the vector field v(x,y) = [u(x,y),v(x,y)] that minimizes the energy functional

$$\varepsilon = \int\int \mu\,(u_x^2 + u_y^2 + v_x^2 + v_y^2) + |\nabla f|^2\,|v - \nabla f|^2\,dxdy \quad (2)$$

The effect of this variational formulation is that the result is made smooth when there is no data.

When the gradient of the edge map is large, it keeps the external field nearly equal to the gradient, but keeps field to be slowly varying in homogeneous regions where the gradient of the edge map is small, i.e., the gradient of an edge map ∇f has vectors point toward the edges, which are normal to the edges at the edges and have magnitudes only in the immediate vicinity of the edges and in homogeneous regions ∇f is nearly zero. μ is a regularization parameter that governs the tradeoff between the first and the second term in the integrand in Eq. 2.

The solution of Eq. 2 can be done using the Calculus of Variations and further by treating u and v as functions of time, solving them as generalized diffusion equations (Xu and Prince, 1998).

DISCRETE COSINE TRANSFORM (DCT) BASED GVF ACTIVE CONTOURS

The transform of an Image yields more insight into the properties of the image. The Discrete Cosine Transform has excellent energy compaction. Hence, the Discrete Cosine Transform promises better description of the image properties. The Discrete Cosine Transform is embedded into the GVF Active Contours.

When the image property description is significantly low, this helps the contour model to give significantly better performance by utilizing the energy compaction property of the DCT.

The 2D DCT is defined as:

$$C(u,v) = \alpha(u)\alpha(v)\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)\cos\left[\frac{(2x+1)u\pi}{2N}\right]\cos\left[\frac{(2x+1)v\pi}{2N}\right] \quad (3)$$

The local contrast of the Image at the given pixel location (k, l) is given by

$$P(k,l) = \frac{\sum_{t=1}^{2(2n+1)-1} w_t E_t}{d_{00}} \quad (4)$$

where,

$$E_t = \frac{\sum_{u+v=t} |d_{u,v}|}{N} \quad (5)$$

and

$$N = \begin{cases} t+1 & t < 2n+1 \\ 2(2n+1) - t & t \geq 2n+1 \end{cases} \quad (6)$$

Here, w_t denotes the weights used to select the DCT coefficients. The local contrast P(k, l) is then used to generate a DCT contrast enhanced image (Jinshan and Acton, 2004), which is then subject to selective segmentation by the energy compact gradient vector flow active contour model using Eq. 2.

RESULTS AND DISCUSSION

The chromosome metaphase image (at 72 pixels per inch resolution) was taken and preprocessed. Insignificant and unnecessary regions in the image were removed, interactively.

Interactive selection of the chromosome of interest was done by selecting a few points around the chromosome that formed the vertices of a polygon. On constructing the perimeter of the polygon, seed points for the initial contour were determined automatically by periodically selecting every third pixel along the perimeter of the polygon.

The GVF deformable curve was then allowed to deform until it converged to the chromosome boundary. The optimum parameters for the deformable curve with respect to the Chromosome images were determined by tabulated studies.

The image was made to undergo minimal preprocessing so as to achieve the goal of boundary mapping in chromosome images with very weak edges.

The DCT based GVF Active contour is governed by the following parameters, namely, σ, μ, α, β and κ. σ determines the Gaussian filtering that is applied to the image to generate the external field.

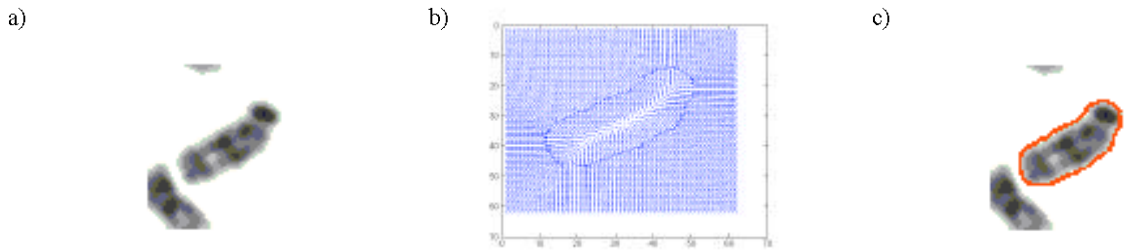

Fig. 10: a) A sample chromosome image, b) DCT based GVF field and c) The segment output image

Larger value of σ will cause the boundaries to become blurry and distorted and can also cause a shift in the boundary location. However, large values of σ are necessary to increase the capture range of the active contour. μ is a regularization parameter in Eq. 2 and requires a higher value in the presence of noise in the image. α determines the tension of the active contour and β determines the rigidity of the contour. The tension keeps the active contour contracted and the rigidity keeps it smooth. α and β may also take on value zero implying that the influence of the respective tension and rigidity terms in the diffusion equation is low. κ is the external force weight that determines the strength of the external field that is applied. The iterations were set suitably.

The DCT based GVF active contours successfully boundary mapped the chromosome images. Figure 10a-c show one example of successful boundary mapping process.

The DCT based GVF active contours have hence been successfully characterized (Britto and Ravindran, 2005c), standardized (Britto and Ravindran, 2005d) and assessed as an efficient tool (Britto and Ravindran, 2006a) for chromosome boundary mapping. Also, the DCT based GVF active contours are also capable of chromosome image classification (Britto and Ravindran, 2006b). We find that the DCT based GVF active contours are able to perform efficiently and are not influenced by variation of the source of the images, namely, the dataset from which they are drawn (Britto and Ravindran, 2006a).

The mathematical formulation for the final contour that approximates the boundary of the chromosome, as seen in Fig. 10c, can be obtained from the coding used for the boundary mapping process and the shape properties of the final contour can be derived from it.

With reference to the chromosomal abnormalities, it can be observed that a suitable template can be used to detect the abnormalities, namely the presence of absence of a particular chromosome.

The shape properties of the final contour, after suitable template formation, in conjunction with the chromosome image classification capability of the DCT based GVF Active Contours (Britto and Ravindran, 2006b), can be used to detect the presence or absence of a particular chromosome. This will yield information regarding the presence of a particular syndrome, among those syndromes discussed earlier.

Classification and Template matching techniques are not discussed here. The main significance of this work consists in establishing that DCT based GVF active contours can be used as a suitable tool for detection of specific chromosomal abnormalities, as it has been found to have good strength for accomplishing this task. Discussion on suitable pattern recognition or template matching techniques have not been given, as any good template matching technique using GVF based DCT active contours for building up the prototype model will succeed. When the prototype template has been built using DCT based GVF active contours, then template matching is an easy task.

Hence the DCT based GVF Active Contours can be used to build up the prototype template model and using suitable template matching technique can be used as a suitable tool for detection of the presence of specific human chromosomal abnormalities.

CONCLUSIONS

The DCT based GVF active contours can be used (in conjunction with appropriate template matching processes) as a suitable tool for detection of specific human chromosomal abnormalities.

ACKNOWLEDGMENTS

The authors extend their heartfelt thanks to Dr. Michael Difilippantonio, Staff Scientist at the Section of Cancer Genomics, Genetics Branch/CCR/NCI/ NIH, Bethesda MD; Prof. Ekaterina Detcheva at the Artificial Intelligence Department, Institute of Mathematics and Informatics, Sofia, Bulgaria; Prof. Ken Castleman and Prof. Qiang Wu, from Advanced Digital Imaging

Research, Texas and Wisconsin State Laboratory of Hygiene-http://worms. zoology.wisc.edu/zooweb/Phelps/karyotype.html for their help in providing chromosome spread images.

REFERENCES

Britto, A.P. and G. Ravindran, 2005a. A review of deformable curves from the perspective of chromosome image segmentation. J. Med. Sci., 5: 363-370

Britto, A.P. and G. Ravindran, 2005b. Comparison of boundary mapping efficiency of gradient vector flow active contours and their variants on chromosome spread images. J. Applied Sci., 5: 1452-1460

Britto, A.P. and G. Ravindran, 2005c. Boundary mapping of Chromosome spread images using optimal set of parameter values in discrete cosine transform based gradient vector flow active contours. J. Applied Sci., (In Press).

Britto, A.P. and G. Ravindran, 2005d. Evaluation of standardization of curve evolution based boundary mapping technique for chromosome spread images. J. Infom. Technol., (In Press).

Britto, A.P. and G. Ravindran, 2006a. Segmentation of chromosome spread images using discrete cosine transform based gradient vector flow active contours-an analysis. J. Med. Sci., 6: 117-124.

Britto, A.P. and G. Ravindran, 2006b. Discrete cosine transform based gradient vector flow active contours-a suitable tool for chromosome image classification. J. Med. Sci., 6: 262-274.

http://www.colorado.edu/MCDB/MCDB2150Fall/notes00/L0001.html (last accessed Sept. 10, 1005).

http://home.earthlink.net/~heinabilene/karyotypes/karyoty.htm (last accessed Sept. 10, 1005).

http://arbl.cvmbs.colostate.edu/hbooks/genetics/medgen/chromo/chromosomes.html (last accessed Sept. 10, 1005).

http://anthro.palomar.edu/abnormal/default.htm.

Jinshan, T. and S.T. Acton, 2004. A DCT based gradient vector flow snake for object boundary detection. 6th IEEE Southwest Symposium on Image Analysis and Interpretation, pp: 157-161.

Kass, M., A. Witkin and D. Terzopoulos, 1987. Snakes: Active contour models. Intl. J. Comp., 1: 321-331.

McInerney, T. and D. Terzopoulos, 1996. Deformable models in medical image analysis. IEEE Proceedings of the Workshop on Mathematical Methods in Biomedical Image Analysis, pp: 171-180.

Prince, J.L. and C. Xu, 1996. A new external force model for snakes. In: 1996 Image and Multidimensional Signal Processing Workshop, pp: 30-31.

www.biology.iupui.edu\biocourses\N100\2K2humancsomaldisorders.html (last accessed Sept. 10, 1005).

Xu, C. and J.L. Prince, 1998. Snakes, shapes and gradient vector flow. IEEE Transactions on Image Processing, 7: 359-369.

Xu, C. and J.L. Prince, 2000. Gradient Vector Flow Deformable Models. In: Handbook of Medical Imaging, Academic Press, pp: 159-169.