

DETECTION OF CHROMOSOME ABNORMALITY USING IMAGE PROCESSING TECHNIQUE FOR HUMAN CHARACTERISTICS

¹Dr.Vanitha.L, ²Aswini.S, ³Gowthami.M, ⁴Hemalatha.N.V

¹Associate Professor, ²Student, ³Student, ⁴Student

¹Department of Electronics and Communication,

¹Prathyusha Engineering College, Chennai, India

Abstract: The proposed system classifies the human chromosomes and detects chromosomal abnormalities automatically without human supervision. Chromosome image is acquired and processed, features are extracted and Linear Vector Quantization classifier is used for classification. LVQ are multiclass classification algorithm by which it improves accuracy and another advantage in LVQ is the algorithm complexity can be adjusted during training. Chromosome abnormality is a missing, extra or irregular portion of chromosomal DNA. There are two types of chromosome abnormality they are numerical abnormality and structural abnormality. Instead of 46 chromosomes there might be 45 or 47 chromosomes is type of abnormality is called numerical abnormality. Structural abnormality occurs when there is a change in the structure or components of a chromosome. In this work only numerical abnormality is detected. The chromosomes are classified according to their length, width, area, entropy, standard deviation and are compared with the normal. This system helps in identifying the chromosome abnormality and detects genetic disorders in infants.

IndexTerms - Matlab, DNA, FIR, IIR, ROI, LVQ, PCA.

I.INTRODUCTION

Chromosomes are several threadlike bodies, it consists of chromatin, which carry the genes. The human being has 23 pairs of chromosome i.e., 46 chromosomes, arranged from 1 to 22 in order of decreasing size and XX and XY for the female and male sex chromosomes respectively. The chromosome consists of protein and a single molecule of DNA that is carried over from parents to young. DNA contains the specific information which makes every living creature unique. A Chromosome band is defined as a section of a chromosome, which shows comparatively darker or lighter stain as compared to the adjoining sections of the identical chromosome. All the twenty four pairs of chromosomes comprise a particular band pattern. The entire chromosome has many segments. The chromosome is typically divided into two arms "short arm" and a "long arm" of the chromosome. The short arm, is the upper half of the chromosome, is identified as the "p arm." The long arm, which is the lower half of the chromosome, is known as the "q arm." In between p arm and q arm there appears a center part called as centromere.

II. METHODOLOGY

BLOCK DIAGRAM

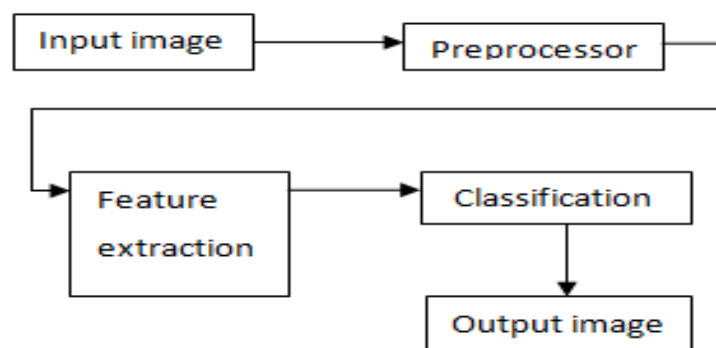


Figure shows the block diagram of classification of chromosome. The main aim of this is the classification of features in an image. This system includes:

- Subsystem to extract chosen features
- Subsystem for classification known as classifier.

The classifier used in the proposed system is LVQ algorithm.

2.1 Modules

In the proposed system there are four modules:

- Pre-processing
- Feature extraction
- Linear vector Quantization classifier for classification
- Detection of chromosome abnormalities

2.2 Pre-Processing

In feature extraction algorithm the noise at the background of images results in poor detection of chromosome feature. Thus pre-processor is used and it removes noise from the image. This is necessary for the consistent extraction of features.

2.3 Feature Extraction

The three main approaches of feature extraction are as follows:

- Statistical approach
- Syntactic or structural approach
- Spectral approach.

Statistical approach is a approach in which a set of statistically extracted features is represented in a vector format in multidimensional feature space. For example: chromosomal abnormality is detected using the characters of chromosome such as length, width, standard deviation, entropy, etc.

Syntactic approach has a texture primitive through which texture is determined, there are a set of placement imperative by which it is planned spatially to produce a complete pattern.

Spectral approach is an approach in which spatial frequencies determine the texture.

On analyzing these three approaches

Statistical approach -It is effective for random patterns

Syntactic approach -More effective

Spectral approach -Less effective

In this paper both syntactical and statistical approach is used for texture based segmentation. Statistical features found by texture-primitive cell are applied for segmentation and classification. Medial axis transformation (MAT) helps in the process of feature extraction by determining the medial line of a chromosome.

The following are the different features extracted from chromosomes:

- (1) Length
- (2) Area
- (3) Entropy
- (4) Standard deviation

(1) Length

The length of every chromosome is determined by including the number of pixels in the medial line. The relative length of the i -th chromosome (l_i) can be obtained by normalizing the medial axis length using the following equation.

$$l_n = \frac{l_i}{l_t}$$

Where l_i ($i = 1, 2, \dots, 24$) is the length of i -th chromosome it is the total length of all 46 Chromosomes of one cell.

(2) Area

The relative area of the i -th chromosome (A_i) can be obtained by counting the pixels of the chromosome body and by normalizing the areas using the following equation

$$A_n = \frac{A_i}{A_t}$$

where A_i ($i = 1, 2, \dots, 24$) is the area of i -th chromosome and A_t is the total area of all 46 chromosomes of one cell.

(3) Entropy

It is a measure of randomness which can be used to characterize the texture of the input image. Entropy is defined as

$$e = \sum(p_i \cdot \log_2(p_i))$$

$$\mu_p = \frac{1}{n^2} \sum_{r=0}^{n-1} \sum_{s=0}^{n-1} P_{r,s}$$

where $p_{r,s}$ is pixel at location (r, s) .

where p contains the histogram counts returned from `imhist`.

(4) Standard deviation

$$e = -\sum_{b=0}^{L-1} p(b) \log_2 p(b) \quad (6)$$

where $p(b) \approx N(b)/n^2$ for $\{0 \leq b \leq L-1\}$,

where L is the number of different values which

$s = \text{std}(X)$, where X is a vector. The output is the square root of a balanced estimator of the variance of the population from which X is drawn, as long as X consists of independent, identically distributed samples.

Chromosome Abnormality

Chromosome abnormalities usually occur due to mistake or damage in the cell division. Chromosome abnormality occurs due to an extra or missing number of chromosomes and they cause birth defects. They are mainly classified into two types numerical abnormality and structural abnormality.

Numerical Abnormality

Numerical abnormality causes abnormality for human species when there is an extra or missing chromosome among 46 chromosome i.e., it may consists of either 45 or 47 chromosome. And this causes birth defect.

Structural Abnormality

Structural abnormality occurs when a part of chromosome is deleted, duplicated or change in chromosome structure. This ultimately leads to change in genetic material which causes birth defects. In this paper we discuss only about the numerical abnormality. It detects whether it consists of extra or less number of chromosome.

III IMPLEMENTATION RESULTS

The segmented chromosome is given as input to the feature extraction algorithm. In which length, area, entropy, standard deviation is determined. And a large data set is collected for training. The collected data set is given as input to the linear vector quantization classifier to classify the chromosome based upon their feature and numerical abnormality is detected.

$$\text{Accuracy} = \frac{\text{Number of correctly classified Chromosomes}}{\text{Total number of Chromosomes}}$$

$$\text{Error Rate} = \frac{\text{Number of misclassified Chromosomes}}{\text{Total Number of Chromosome}}$$

Tools for Experiment

MATLAB is used as a programming language for feature extraction, chromosome classification and detection of chromosomal abnormalities.

IV CONCLUSION

The proposed approach achieved 94% classification accuracy on data set, which is very promising. Based on the classification, the system is capable to detect the numerical abnormalities of the chromosomes. The chromosome images are filtered using the images preprocessing technique and features are extracted and LVQ classifier.

V FUTURE ENHANCEMENT

This work deals only with the analysis of numerical abnormality of chromosomes by means of LVQ classifier. This work can be extended to diagnose different structural abnormality of chromosomes. The efficiency of classification can be increased by using efficient classifier.

VI. REFERENCES

- [1] Akila Subasinghe, "An image processing algorithm for accurate extraction of the centerline from Human Metaphase Chromosome", pp. 5652017, 2010.
- [2] Anu, "An automated detection and morphological classification of numerical abnormalities in human chromosomes", IEEE, pp.0337, 2015.
- [3] Carlos Eduardo, "Adapting learning parameter transition in the generalized learning vector quantization family of classifier", IEEE, pp.5594950, 2010.
- [4] Manish Mishra, "A view of Artificial Neural Network", IEEE, pp.701275, 2015.
- [5] Mousami V. Munot, Madhuri A. Joshi, "Automated Karyotyping of Metaphase Cells with Touching Chromosomes", pp.3875175, IEEE, 2012.
- [6] Ravi Uttamatin, "Band Classifications based on chromosome shapes", IEEE Conference, pp.6687672, 2013.
- [7] Ravi Uttamatin, "Chromosome classification for metaphase selection", IEEE, pp.6645903, 2013
- [8] Saranya, Loganathan, Ramaprabha, "Efficient feature extraction and classification of chromosomes", pp.7396086, 2016.
- [9] Shervin Minaee, "A geometric approach to fully automatic chromosome segmentation", IEEE, pp.1112.4164, 2014.
- [10] Tanvi, Renu Dhir, "An Efficient Segmentation Method for overlapping chromosome images" pp.3894861, IJCA, 2014.