A Database System for the Management and Karyotype Analysis of Chromosome Image Data

Akihiro Kawaguchi   **Kenji Satou**

akawa@jaist.ac.jp   ken@jaist.ac.jp

School of Knowledge Science, Japan Advanced Institute of Science and Technology,
1-1 Asahidai, Tatsunokuchi, Ishikawa 923-1292, Japan

1   Introduction

In many laboratories of biology which usually doing FISH and GISH experiments, it is important to store, retrieve, and manage large amount of chromosome image data easily. For this purpose, conservative way of data management, that is, each technician keeps their experimental data in their own PCs, will ruin for thousands or millions of image data. Furthermore, version management of image data is also needed in such laboratories since karyotype analysis means a series of image processing from raw image data taken by cooled CCD camera to final image data in which all the chromosomes are cropped, directed vertically, and sorted in the order of their size. To assist such a chromosome image data management, we are developing a database system for storing, retrieval, and karyotype analysis of them (Fig. 1).
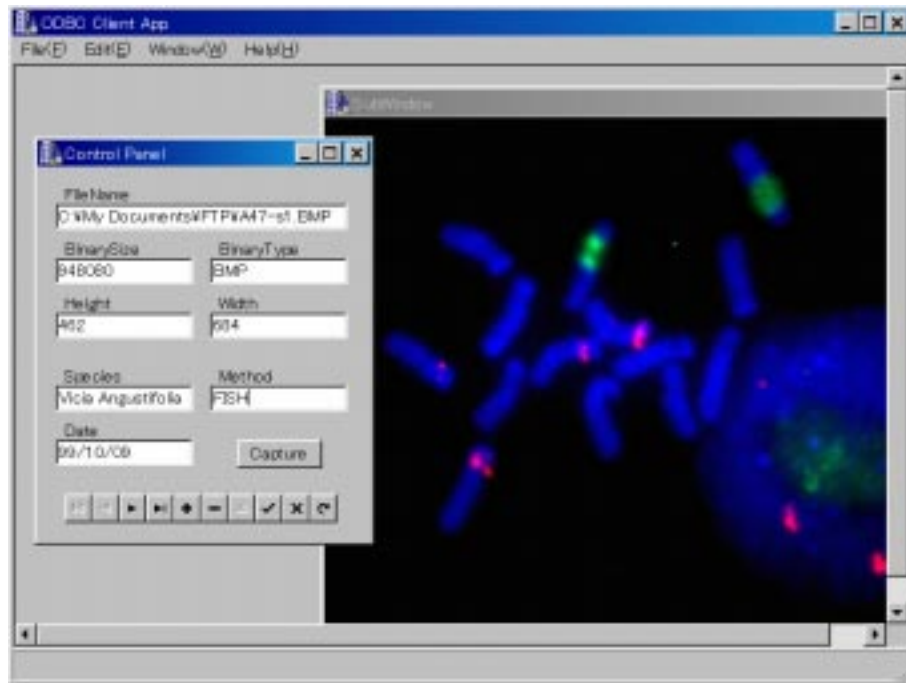

Figure 1: A screenshot of client application.

2   Features of the System

In this section, we describe features of our system.

2.1   System Design

For allowing easy operation from technician's PC, this system consists of Windows client and back-end DBMS. Since the client is programmed to connect backend DBMS via ODBC (Open Database Connectivity), it can be freely chosen from any DBMSs (e.g. ORACLE, Sybase, etc.) as long as there exist ODBC drivers for them. Currently, PostgreSQL [4] is adopted in our system.

2.2   Method of Information Retrieval

The system uses two search methods. One is a function of DBMS, another is in client side. DBMS search can respond to user's request by SQL, that uses search key about date, species, experimental methods, and so on. On the other hand, client side search function enables retrieval of similar images. The client performs feature extraction from colors and shapes of each chromosome in a image [2, 3]. Since these two methods are integrated in the user interface, a user can be unconscious about the difference of retrieval methods.

2.3   Version Management of Image Processing

In the analysis of FISH and GISH experimental data, usually it is needed to reprocess the same raw image data for analyzing them from various viewpoints. So the system should be able to keep parent-child relationships on the transformation process of image data for allowing easy navigation and comparison of each experimental data. Using this relationship as a search key, a user can retrieve a stack of images derived from the same raw image data.

2.4   Automated Karyotype analysis

Karyotype analysis is one of the major methods in chromosome image analysis [1]. Today, in many laboratories without excellent and expensive software for karyotype analysis, it is manually done by using photoretouch software. Therefore, we are trying to implement automated karyotype analysis and integrate the function into our system.

Acknowledgments

References

[1] Kato, S., Hirose, T., Akiyama, Y., O'Neill, C. M., and Fukui, K., Manual on the chromosome image analyzing system III, CHIAS III, *Research Reports on Agricultural Development in the Hokuriku Area*, 36, 1997.

[2] Hiroike, A., Mori, Y., and Sakurai, A., Shape discrimination by using integral features, *The Transactions of the Institute of Electronics, Information and Communication Engineers. D-II*, J80-D-II, 1:81–91, 1997.

[3] Matsuzawa, Y., Region extraction by competition of multiple active contour models, *School of information science, JAIST, master's thesis*, 1999.

[4] PostgreSQL Official Web Page, http://www.postgresql.org/