Development of a Spot Matching Module in Image Analysis System DDGEL for 2D Gel Electrophoresis

Tatsuya Akutsu [1]
takutsu@ims.u-tokyo.ac.jp

Akira Ohyama [2]
akr@hydra.mki.co.jp

Kyotetsu Kanaya [2]
kanaya@hydra.mki.co.jp

Asao Fujiyama [3]
afujiyam@lab.nig.ac.jp

[1] Human Genome Center, Institute of Medical Science, University of Tokyo
4-6-1 Shirokanedai, Minato-ku, Tokyo 108, Japan

[2] Department of Bioscience Systems, Mitsui Knowledge Industry Co., Ltd.
2-7-14 Higashinakano, Nakano-ku, Tokyo 164, Japan

[3] National Institute of Genetics, 1111 Yata, Mishima-city, Shizuoka 411, Japan

Abstract

We have been developing an image analysis system named DDGEL for 2D gel electrophoresis of genomic DNA. Recently, we have developed a program module for finding a correspondence of spots between two gel electrophoresis images.

1 Introduction

Although hybridization-based methods and PCR-based methods have been successfully applied to analysis of genomic DNA, these methods are not suited for rapid analysis of the entire genomic DNA. Recently, a method called RLGS (Restriction Landmark Genomic Scanning) has been developed in order to detect and analyze the genetic alterations by observing the entire genomic DNA after separating DNA fragments in a single two-dimensional slab gel [1]. To analyze gel images obtained by the RLGS method, a lot of tasks must be done: thousands of spots must be detected where each spot corresponds to a gene; a correspondence of spots between two images must be detected; links between spots and genetic information must be classified and stored in a database. We have been developing a software tool named **DDGEL** so that such tasks can be done automatically or semi-automatically [2]. In this short abstract, we give a brief description about a recently developed pattern matching program for detecting a correspondence of spots between two images.

2 Method and Result

Spots are detected using standard image processing techniques. For two sets A and B of spots, we find a correspondence of spots between two sets by the following procedure.

(1) A few pairs of matching spots (*landmarks*) are detected (manually or automatically).

(2) In order to adjust distorsion, B is transformed using a matrix derived by applying the least squares fitting method to landmarks.

(3) A is divided into several rectangular regions.

(4) For each rectangular region, a correspondence between spots in this region and spots in B is computed using bipartite graph matching. Moreover, this correspondence is refined via iterative improvement using least squares fitting and bipartite graph matching.

(5) A correspondence between the entire regions is refined using bipartite graph matching.

An example of a result of spot matching applied to Human DNA is shown in Figure 1. For images with $\approx 1,150$ spots, matching was obtained in 12 seconds using SUN Ultra-2 workstation (300MHz). Even for a case that distortion between two images was not small, there was a little difference between the matching detected by the system and that detected by a human expert.

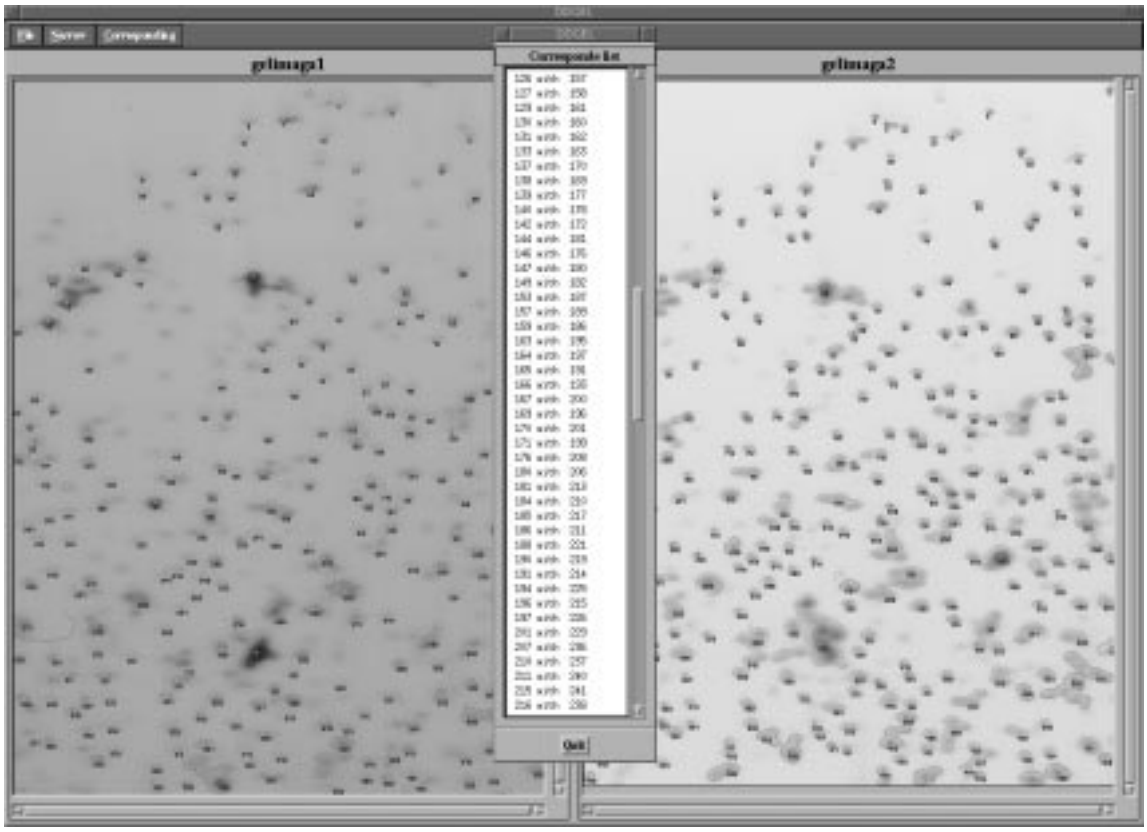

Figure 1: Example of spot matching, where parts of images are shown here. Matched spots are listed in the middle window and unmatched spots are highlighted with green edges.

3 Conclusion

Owing to the development of the spot matching module, DDGEL system has become a powerful tool for 2D gel image analysis. DDGEL system runs on a Unix workstation with SunOS 4.1.2, Solaris 2.3, or an upper-compatible operating system, and program codes with annotated standard gel data of Human DNA are now available via FTP (see `http://www.hgc.ims.u-tokyo.ac.jp`).

References

[1] Hatada, I. *et al.*, "A genomic scanning method for higher order organisms using restriction sites as landmarks," *Proc. Natl. Acad. Sci. USA.*, 88:9523–9527, 1991.

[2] Ohyama, A., Akutsu, T. and Fujiyama, A., "A Software Tool for Mapping Human Genome by Chromosome-specific Two- Dimensional Electrophoresis Method," *Proc. Genome Informatics Workshop 1995*, pp. 144-145, 1995.