

cDNA 마이크로어레이 이미지를 위한 그래프 모델과 분석 알고리즘

(A Graph Model and Analysis Algorithm for cDNA
Microarray Image)

정 호 열[†] 황 미 념[†] 유 영 중^{**} 조 환 규^{***}

(Ho-Youl Jung) (Mi-Nyeong Hwang) (Young-Jung Yu) (Hwan-Gue Cho)

요 약 본 논문에서는 마이크로어레이 처리를 위한 새로운 이미지 분석 알고리즘과 격자 반점들의 위상 정보를 이용하여 격자의 위치를 결정하는 방법을 제시한다. 마이크로어레이는 유전자의 발현량의 측정을 위해서 수만 혹은 수십 만개의 유전자에 대해서 한번에 실험을 할 수 있는 장비이다. 한번의 실험으로 생성되는 데이터 양이 엄청나게 많기 때문에 자동화된 분석이 필요하다. 이 마이크로어레이의 실험 결과는 16-비트 회색조 이미지 파일로 하나의 유전자가 여러 개의 픽셀로 뭉쳐져 있는 반점(spot)이 격자 구조 형태로 나타난다. 본 논문에서 이미지 데이터에서 그래프 구조를 생성하여 이들 반점이 어느 격자에 속하는지 결정하는 알고리즘과 격자 구조의 기울어짐을 측정하여 격자의 정확한 위치와 모양을 결정하는 방법을 제시하고 실제 이미지 데이터를 통한 많은 실험 결과를 보여 준다.

키워드 : 이미지 처리, 그래프 모델, DNA 칩

Abstract In this paper we propose a new image analysis algorithm for microarray processing and a method to locate the position of the grid cell using the topology of the grid spots. Microarray is a device which enables a parallel experiment of 10 to 100 thousands of test genes in order to measure the gene expression. Because of the huge data obtained by a experiment automated image analysis is needed. The final output of this microarray experiment is a set of 16-bit gray level image files which consist of grid-structured spots. In this paper we propose one algorithm which locates the address of spots (spot indices) using graph structure from image data and a method which determines the precise location and shape of each spot by measuring the inclination of grid structure. Several experiments are given from real data sets.

Key word : Image processing, graph model, DNA chip

1. 마이크로어레이 개요

마이크로어레이는 어떤 슬라이드 박편에 여러 가지 종류의 DNA 염기가닥(sample clone)을 행렬 형태로 심어 놓은 것이다. 이 마이크로어레이를 이용하면 이전

의 단순한 실험에서 얻을 수 있는 것 이상의 매우 많은 유전자에 관한 실험을 동시에 할 수 있다. 아래 그림 1에 나타나 있듯이 DNA 염기가닥은 여러 가지 장기기관이나 개체에서 추출되는데 그 양이 매우 적으므로 이것은 PCR(Polymerase Chain Reaction)이라는 특별한 방법을 통하여 증폭하여 많은 양의 같은 유전정보를 가진 개체를 만들어 낸다. 그리고 이것은 자동화된 로봇 프린팅 기술을 이용하여 약 1인치 정도의 면적에 대략 100×100개의 반점을 찍는다. 이러한 작업을 하는 기계를 보통 마이크로어레이라고 부른다[1]. 그리고 이러한 수만 개의 유전자들이 붙어있는 슬라이드를 보통 DNA 칩 혹은 마이크로어레이라고 부른다. 그런데 유전자의 종류에 따라서 올리고 칩(Oligonucleotide chip)이

· 이 논문은 2000년도 한국진흥재단의 지원에 의하여 연구되었음
(KRF-2000-041-E00304)

† 비 회 원 : 부산대학교 전자계산학과

hyjung@pearl.cs.pusan.ac.kr

mnhwang@pearl.cs.pusan.ac.kr

** 비 회 원 : 부산외국어대학교 컴퓨터공학과

yjyu@taejo.pufs.ac.kr

*** 정 회 원 : 부산대학교 전기전자정보컴퓨터공학부 교수

hgcho@hywon.pusan.ac.kr

논문접수 : 2001년 11월 15일

심사완료 : 2002년 4월 4일

라는 것도 있다. 올리고 칩은 올리고당이라는 분자단위 단위로 DNA 염기서열을 이루는 아데닌(A), 구아닌(G), 티민(T), 시토신(C)을 직접 만들어 붙인 것이다. 예를 들어 올리고당을 이용하면 ATTGCGA..GT등과 같은 40 base정도의 염기서열을 만들 수 있다. 올리고 칩을 만드는 것은 화학적인 방법이 아니라 일반적인 반도체를 만들 때 사용하는 식의 리소그라피를 사용하므로 그 직접도를 더 높일 수 있다. 본 논문에서는 cDNA를 유전자로 사용하여 슬라이드로 심어둔 칩으로 가장 보편적으로 사용되는 칩에 의해 생성된 이미지를 기본으로 한다.

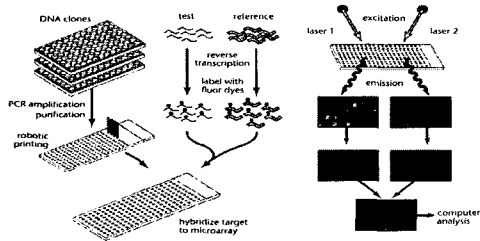


그림 1 DNA chip의 제작과정과 실험과정[1].

cDNA는 complement DNA라고 불리는 특별한 DNA로서 병의 진단이나 유전정보를 탐색하는데 매우 중요하게 사용되는 DNA이다. cDNA는 일반적인 염기서열에서 단백질을 만들어내는 DNA 부서열(sub sequence : coding region)들을 mRNA가 전사(transcription)를 수행하는데 이 mRNA에 역으로 전사되어 그 coding 부서열의 DNA정보를 역전사한 DNA 서열이다. 위의 그림 1의 예와 같이 각 DNA 칩의 반점에는 여러 유전인자가 붙어있는데 우리는 어떤 특정한 유전인자와 반응하는 인자를 검사할 수 있다. 이 검사는 위의 그림과 같이 어떤 유전자의 한 끝에 특별한 색을 내는 형광 물질을 붙인다 [2].

이러한 반응이 끝난 뒤에 그 DNA 칩에 특별한 레이저를 비추면 이전에 반응시킨 테스트 유전자와 참조 유전자가 각 반점 샘플 유전자에 반응한 정도에 비례하여 형광물질도 축적되어 있으므로 바로 그 양에 따라서 다른 양의 형광빛이 나타나게 된다. 따라서 그 형광의 양(Cy-3, Cy-5)이 바로 테스트와 참조 유전자가 얼마나 샘플 유전자에 반응하였는지를 나타내는 지표가 된다.

이렇게 DNA 칩의 각 반점에 올라가 있는 형광물질의 양을 자동적으로 스캐닝하여 이를 16-bit 이미지 파일로 만들어주는 장치를 보통 DNA 칩 스캐너라고 한

다. 결국 최종의 실험결과는 그림 2에서처럼 각 샘플 유전자에 얼마나 잘 테스트 유전자와 반응하는지가 바로 형광의 정도로 나타나게 된다. 이것을 조사하여 테스트 유전자와 잘 반응하는 유전자, 그렇지 않은 유전자 등으로 구분할 수 있다.

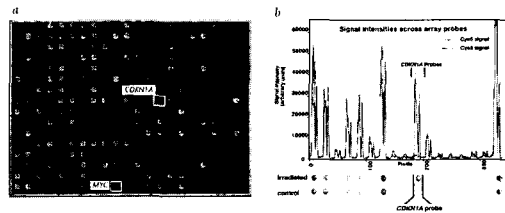


그림 2 DNA 칩 스캐너를 통하여 얻은 이미지와 어떤 한 행에 있는 각 반점들에서 적색 반응을 y축으로 그린 그래프[1].

본 논문에서는 마이크로어레이 이미지를 분석하는 방법에 있어서 각 반점들의 격자 상의 위치를 결정하여 주는 알고리즘과 이미지가 전체적으로 기울어진 경우에 정확한 위치와 격자의 기울어진 모양을 계산하는 방법에 관하여 설명한다. 이를 위해서 이 마이크로어레이 이미지를 그래프 표현 방법으로 변형을 하여 전산학적인 측면에서 자동화된 격자 주소 결정 방법을 설명하려고 한다. 이를 위해 이미지의 그래프 표현 방법에 대해서 알아보고 다음으로 이 그래프 구조에 기반한 격자 주소 결정 방법에 관하여 설명한 후 기울어진 이미지에 대한 정확한 격자 모양과 위치를 결정하는 방법에 관하여 알아본다.

2. 격자 주소 결정 문제

마이크로어레이 이미지에 대한 분석에서 최종적으로 구하고자하는 것은 각 유전자의 발현 정도 즉 이미지 상에서는 각 유전자가 놓여져 있는 위치에 해당하는 격자 내부 픽셀의 강도이다. 그런데 격자 내부 픽셀의 강도를 측정하려고 하면 이미지상의 어느 영역이 어느 격자 내부에 해당하는지를 결정하여 주어야 한다. 여기서 문제가 되는 것은

1) 해당되는 특정 영역을 어떻게 계산할 것인가

2) 특정 영역의 색의 강도는 어떻게 측정할 것인가

의 2가지를 이야기 할 수 있다. 전자는 일반적으로 격자 결정(grid fitting) 또는 격자의 인택스를 부여한다는 의미로 주소 결정(addressing)이라 하고 후자는 반점 결정(spot fitting)이라고 한다. Brändle등은 전자의 문제

가 해결되어 있다는 가정 하에서, 즉 격자 결정이 완료된 상태에서 반점 결정 시에 발생하는 문제점들에 관한 해결 방법에 관해서만 논하였다[3]. 물론 반점 결정도 중요한 문제이지만 격자 결정이 정확히 이루어져 있지 않은 상태에서는 정확한 색의 강도를 측정한다는 것은 불가능하므로 격자 결정의 중요성을 간과할 수 없는 것이다. Buhler 등은 이미지상의 반점 위치를 찾으면서 이들 반점의 모양과 위치에 대한 품질 제어 및 측정에 대한 방법론과 격자 주소 결정을 위한 방법으로는 여러 실험 이미지에 대해서 평균적인 반점의 위치를 미리 측정하고 마이크로어레이의 물리적인 정보를 이용하여 각 반점들의 격자 주소를 결정하는 방법을 제시하였다[4]. 하지만 이런 방식의 격자 주소 결정은 이기종의 마이크로어레이 실험에 의한 이미지 분석은 불가능하다는 단점을 가지고 있다. Brown 등은 마이크로어레이 이미지 데이터를 분석하는 데 있어서 고려되어야 할 품질 제어를 측정하기 위해 통계적인 측정 방법을 사용하였다[5]. 그리고 Steinfath 등은 마이크로어레이에 의해서 생성되는 이미지에 대한 분석과 이미지 분석의 결과인 유전자 발현 프로파일을 클러스터링한 결과를 보여주는 마이크로어레이 분석 도구를 소개하였다[6]. Steinfath 등은 이미지 데이터 분석에 있어서 유전자가 붙어 있는 부영역이 한 개만 존재하는 경우에는 완전 자동화 분석 방법을 사용하고 있고 기울어져 있거나 왜곡 현상이 있는 이미지 데이터에 대해서 보정하는 방법을 제시하고 있다. 하지만 분석하고자 하는 이미지 데이터에서 유전자가 붙어 있는 부영역이 2개 이상인 경우에는 사용자의 개입이 필요로 하며 기울어져 있는 이미지에 대한 보정을 부영역의 4개의 구석 점을 연결하는 간선들의 평균 기울기를 측정하므로 상당히 정확한 마이크로어레이에서는 사용 가능하겠으나 스팟팅이 부정확하여 위치 오류율(positional error rate)이 높은 마이크로어레이에서는 사용이 불가능하다.

격자 구조란 행렬 모양으로 영역이 구분되어 있는 것을 말한다. 격자 모양으로 구분하는 것은 격자 내부(그리드 셀, grid cell)에 존재하는 데이터를 각각 구분하여 분석하기 용이하도록 하기 위함이다. 이런 분석의 용이함으로 격자 구조는 마이크로어레이와 같은 이미지 데이터에서 사용되고 있고, n -well 플레이트(plate)를 CCD(Charge Coupled Device) 카메라로 직접 촬영한 이미지 데이터도 마찬가지로 구조를 가지고 있다.

마이크로어레이 이미지는 여러 개의 반점들이 행렬 모양으로 놓여져 있는 것을 부격자(subgrid)라고 하는데 이들 부격자들도 한 장의 이미지에 행렬 형태로 놓여있

는 구조를 나타내고 있다. 이런 마이크로어레이 이미지의 이상적인 상태에서의 조건은 다음과 같은 것들이 있다:

- 1) 모든 부격자들의 크기는 같다. 즉, 하나의 부격자의 행과 열에 존재하는 반점들의 개수는 각각 동일하다.
- 2) 부격자 사이의 간격은 일정하다.
- 3) 반점의 위치는 반점이 속한 행과 열의 직선의 교점 위에 존재한다.
- 4) 반점의 크기와 모양은 완벽한 원형이고 모든 반점들에 대해서 동일하다.
- 5) 격자의 위치는 실험에 사용한 모든 슬라이드들에 대한 이미지들 안에서는 동일하다.

하지만 실제 실험 결과 이미지에서는 이러한 이상적인 조건들을 만족하는 경우는 현재의 기술 수준으로는 매우 드물다. 따라서 이런 이유로 인해 분석에 대한 사용자의 개입이 필요로 하고 이런 사용자의 개입 정도에 따라 마이크로어레이 분석 시스템을 세 가지의 범주로 구분할 수 있다.

- 1) 보조 소프트웨어 수준(Assistance software) : 실험자에게 단순히 전산적 처리를 위한 사용자 그래픽 인터페이스만 제공하여 주는 수준이다.
- 2) 부분적 자동 처리 시스템(Semi-automated processing system) : 시스템이 먼저 각 반점들의 위치와 모양을 결정하여 주면 실험자가 직접 하나씩 살펴서 잘못 결정된 반점들에 수정을 하여 주는 시스템이다.
- 3) 완전 자동 처리 시스템(Fully automated processing system) : 실험자의 최소한의 개입으로 수행되는 시스템으로 일반적으로 부격자의 크기와 부격자내의 반점들의 개수만 입력하여 주면 나머지 모든 과정을 시스템이 완전 자동으로 처리하여 주는 시스템으로 분석에 소요되는 비용을 획기적으로 절감시켜 주는 방식이다.

격자 구조를 가진 이미지 데이터를 분석하는 많은 시스템들이 현재 나와있는데 사용자가 격자 구조를 지정하고 반점의 위치도 직접 결정하는 수작업이 많이 들어가는 시스템에서부터 사용자의 수작업을 최대한 줄여 시스템이 반점의 위치를 결정하여 주는 자동화되어진 시스템까지 매우 다양하다. 부분 자동화 시스템의 한 예로는 ScanAlyze[7, 8]와 같은 이미지 분석 시스템이 있는데, 사용자가 각 반점들의 위치를 이미지를 보면서 대충 맞추어 주면 시스템이 정확한 위치로 재 수정하여 주는 방식을 사용하고 있다. 그리고 BioDiscovery사의 AutoGene [9]이나 Media Cybernetics사의 Array-Pro

[10]같은 분석 시스템은 사용자의 간단한 입력으로 모든 반점의 위치를 자동으로 찾아주는 완전 자동화된 시스템이다.

이전의 전통적인 실험에서는 결과로 생성되는 데이터의 양이 그렇게 많지 않기 때문에 실험에 소요되는 시간과 분석에 소요되는 시간이 많은 차이가 나지는 않았으나 현재의 마이크로어레이 기술에 의하면 한번에 많은 양의 유전자들을 동시에 처리할 수 있기에 실험 시간 대비 분석 시간의 비가 거의 1:100 정도를 보인다 [11]. 따라서 분석에 드는 비용을 줄이기 위해 완전 자동화된 분석 시스템이 필요하다. 완전 자동화의 걸림돌로 작용되는 부분은 이미지 분석에 있어 이미지 데이터의 이상적인 조건을 만족시키지 못함에 의해 주소 할당의 자동화가 어렵다. 따라서 이 과정에서 사용자의 수작업이 많이 필요로 하고 분석에 소요되는 시간과 비용도 커진다. 따라서 자동화된 주소 결정 알고리즘의 필요성이 대두되는 부분이다.

3. 격자 주소 결정 알고리즘

본 논문에서는 격자 주소 결정의 자동화를 위하여 마이크로어레이 이미지를 그래프 구조로 변경한 후에 처리를 한다. 입력으로 주어진 이미지에서 어떤 임계값, T , 이상의 연결 요소(connected component)를 먼저 찾고, 이들 연결 요소에 대한 크기 필터(size filter)를 적용한 후에 남은 연결 요소들의 무게 중심을 생성하고자 하는 그래프의 정점으로 한다. 다음은 그래프를 구성하는 방법과 생성된 그래프를 기반으로 격자 주소 결정 알고리즘에 관하여 설명한다.

3.1 마이크로어레이 이미지의 그래프 표현

본 논문에서는 격자 주소 결정 알고리즘을 위해 기하적 표현 방법인 그래프에 의한 표현 방법을 제시한다. 제시하는 그래프는 하나의 정점이 최대 4개의 간선을 가지는 것으로, 입실론 그래프, $G_\epsilon = (V, E)$ 라고 부르기로 한다. $G_\epsilon = (V, E)$ 는 방향 그래프(directed graph)로서 정점의 집합 V 와 간선의 집합 E 로 구성되어 있다. 그리고 $\forall u \in V$ 는 최대 4개의 간선을 가지고 이들 4개의 간선은 상하좌우의 방향을 나타낸다. 여기서 다음의 조건을 만족하는 u 의 상하좌우 정점의 집합을 각각 V^u, V^d, V^l, V^r 이라고 하면 $|u.x - v.x| < \epsilon$ 또는 $|u.y - v.y| < \epsilon$, 여기서 $v \in V^u, V^d, V^l, V^r$ 이고 이들 집합 중에서 실제 $G_\epsilon = (V, E)$ 의 간선이 되는 u 의 이웃한 정점을 각각, u^u, u^d, u^l, u^r 이라고 하면

$$\begin{aligned} u^u &= \min\{v \in V^u \mid |e(u, v)|\} \\ u^d &= \min\{v \in V^d \mid |e(u, v)|\} \\ u^l &= \min\{v \in V^l \mid |e(u, v)|\} \\ u^r &= \min\{v \in V^r \mid |e(u, v)|\} \end{aligned}$$

이다. 여기서 $|e(u, v)|$ 는 간선의 양 정점 u 와 v 사이의 거리를 의미한다.

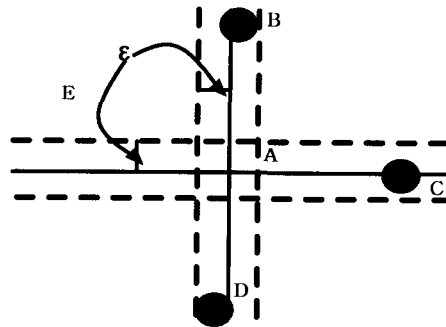
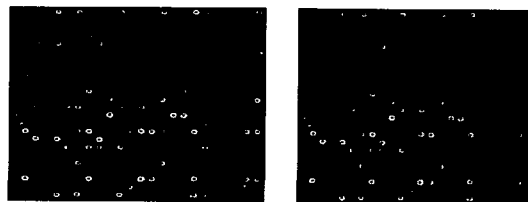


그림 3 $G_\epsilon = (V, E)$ 를 구성하는 과정 : 반점 A에 대한 이웃 반점들의 연결성.

그림 3에서는 반점 A에 대한 4개의 이웃 반점들에 대한 간선 연결성에 대해서 보여 주고 있다. 반점 A에 이웃한 4개의 반점 중에서 B, C, D의 중심은 A의 중심에서 x축 또는 y축 방향으로 ϵ 범위 내에 존재하는 반점들 중에서 각각 상, 우, 하 방향으로 가장 가깝기 때문에 간선으로 연결되고 E의 중심은 ϵ 범위 밖에 존재하기 때문에 좌의 방향으로 가장 가까운 반점이라 할지라도 연결되지 않는다. 그리고 $G_\epsilon = (V, E)$ 는 격자 구조의 특징을 가지고 있다. 즉, 하나의 반점이 가질 수 있는 간선이 4개로 각각은 상하좌우의 방향성을 가지고 있으므로 완벽한 격자 구조는 아니라 할 지라도 부분적 격자 구조의 특성을 그대로 가지고 있어 뒤에 격자 주소 결정에 유용하다.



(a) 마이크로어레이 이미지 (b) $G_\epsilon = (V, E)$

그림 4 마이크로어레이 이미지를 $G_\epsilon = (V, E)$ 로 표현한 그림

$G_\epsilon=(V, E)$ 에서 정점 $v \in V$ 는 이미지 처리 단계에서 연결 요소의 중심을 사용하여 계산을 하는데 연결 요소의 중심을 결정하는 방식은 각 연결 요소의 MBR (Minimum Bounding Rectangle)의 중심을 이용하는 것과 무게 중심(center of mass)을 사용하는 것의 두 가지 방법이 있으나 실험에 의하면 마이크로어레이에서 각 반점들이 원형에 가깝다는 특징으로 큰 차이를 보이지는 않았다. 그림 4는 주어진 마이크로어레이 이미지의 일부인 부영역을 $G_\epsilon=(V, E)$ 로 표현한 그림이다.

3.2 격자 주소 결정 알고리즘

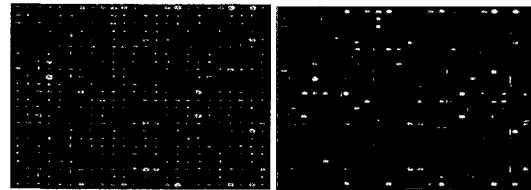
이전까지 마이크로어레이 이미지가 주어지면 이를 그래프를 이용해서 표현하는 방법에 대해서 알아보았다. 이번에는 생성된 그래프를 이용하여 이미지상의 반점들이 어느 격자 주소에 해당되는지를 결정하는 격자 주소 결정(grid fitting 혹은 addressing)에 대해서 설명한다. 이미지상의 특정 영역이 격자 구조의 어느 주소에 해당하는지를 결정하는 문제가 마이크로어레이 이미지에서 왜 어려운 작업인지는 이미 설명을 하였다. 본 논문에서 우리는 사용자의 개입을 최소로 하는 격자 주소 결정 알고리즘을 제시한다.

마이크로어레이 실험에서는 랜드마크(landmark)라는 것을 두는데 이는 스캔한 이미지에서 부영역이 어느 부분에 해당하는지를 표시하는 수단으로 사용되고 있다. 우리는 이런 랜드마크와 비슷한 방식으로 만약에 $G_\epsilon=(V, E)$ 의 정점 V 중에서 어떤 정점의 격자 주소 값을 알 수 있다면 $G_\epsilon=(V, E)$ 의 나머지 모든 정점들에 대한 격자 주소를 그래프 검색 알고리즘을 응용하여 계산할 수 있다. 그리고 $G_\epsilon=(V, E)$ 의 정점들은 강 발현 반점($I_{i,j} > T$ 이고 $I_{i,j}$ 는 마이크로어레이 이미지 I 의 (i, j) 픽셀 강도)들로 이루어져 있으므로 약 발현 반점에 대한 격자 주소 결정은 아직 하지 않은 상태이다. 하지만 이웃한 강 발현 반점의 격자 주소를 기반으로 사용자 입력에 의한 반점과 반점사이의 평균 거리, d_μ 를 사용하여 계산할 수 있다. 기본적인 격자 주소 결정 알고리즘은 다음과 같다:

```

Algorithm: Grid indexing with landmarks
Input:  $V$  in  $G_\epsilon=(V, E)$ 
Output: Grid index,  $C_i$  corresponding to  $V$ 
 $V_p=V$ ; /* the set of not-indexed spots */
 $V_a=\hat{v}$ ; /* the set of indexed spots */
do
  Find  $e(u, v)=\min\{e(u, v) \in E, u \in V_p, v \in V_a \mid e(u, v)\}$ ;
  Remove  $u$  from  $V_p$ ; add  $u$  into  $V_a$  ;
  Compute grid index for  $u$  based on grid index of  $v$ ;
while  $|V_p| \neq 0$ 
    
```

여기서 V_p 는 아직 격자 주소를 결정하지 않은 정점들의 집합을 의미하고 V_a 는 격자 주소가 완료된 정점들의 집합을 각각 의미한다. 그리고 \hat{v} 는 격자 주소를 알고있는 랜드마크 정점(landmark vertex)을 의미한다. 상기의 격자 주소 결정 알고리즘은 그래프 탐색 알고리즘을 기반으로 하고 있다. 방문하는 정점, u 와 연결되어진 다른 한쪽의 정점, v 의 격자 주소를 (i, j) 라고 하면 u 의 격자 주소, $(i+a, j+b)$ 는 두 점간의 방향성 즉 상, 하, 좌, 우와 반점 사이의 평균 거리, d_μ 를 고려하여 a 와 b 의 값을 결정한다. 여기서 u 에 대해서 v 가 상에 위치하면 $a=-[d(u, v)/d_\mu+r_\mu]$, $b=0$ 이고, 하에 위치하면 $a=[d(u, v)/d_\mu+r_\mu]$, $b=0$, 좌에 위치하면 $a=0$, $b=-[d(u, v)/d_\mu+r_\mu]$, 우에 위치하면 $a=0$, $b=[d(u, v)/d_\mu+r_\mu]$ 이다. 여기서 $d(u, v)$ 는 두 정점 u 와 v 사이의 직선 거리를 의미하며 r_μ 는 사용자 입력에 의한 마이크로어레이의 반점들의 평균 반지름을 나타낸다.



(a) $G_\epsilon=(V, E)$ (b) 격자 주소가 할당된 결과

그림 5 그래프 탐색 알고리즘에 기반한 격자 주소 할당

이와 같은 격자 주소 결정 방식은 그래프 탐색 알고리즘을 기반으로 하고 있지만 격자 주소가 결정되지 않은 반점을 선택하는 단계에서 간선의 길이가 가장 짧은 정점을 선택하기 때문에 전체적인 오차를 줄여 정확한 주소 결정을 가능하도록 한다. 그리고 $G_\epsilon=(V, E)$ 의 정점들, 즉 강 발현 반점들에 대한 격자 주소 결정이 완료되면 나머지 비어 있는 격자 주소, 즉 약 발현 반점들에 대한 격자 주소 결정을 하여 주는데 이것은 이미 주소가 결정되어 있는 가장 가까운 거리에 존재하는 이웃한 강 발현 반점들을 찾아 상하좌우의 관계를 고려하여 격자 주소를 결정할 수 있다. 그림 5는 그래프 탐색 알고리즘에 기반한 격자 주소 할당의 결과를 보여 주고 있다.

격자의 크기 $M=r \times c$ 에서, r 은 격자의 행의 개수 그리고 c 는 격자의 열의 개수라고 하면, 일반적으로 마이크로어레이 이미지에서 강 발현 반점의 개수는 $M/3$ 이

상이고 격자 구조 상에서 한쪽 방향으로 편향되어 나타나지 않기 때문에 전체 반점들에 대한 격자 주소 영역을 안전하게 할당 할 수 있다. 이런 경우에는 우리가 제시한 알고리즘은 한번에 하나의 정점씩 할당하고 있는데 할당되는 정점을 결정할 때 다음의 식을 사용하는데

$$e(u, v) = \min_{\substack{e(u, v) \in E, \\ u \in V_p, v \in V_a \setminus \{e(u, v)\}}} |e(u, v)|$$

여기서 최소 값을 가진 정점을 계산하는데 필요한 시간은 $|V_a|$ 와 비례한다. 그리고 하나의 정점은 최대 4개의 이웃한 정점을 가지고 있으므로 제시한 알고리즘의 시간 복잡도는

M 이 짝수일 때는

$$\begin{aligned} & 4 \cdot 1 + 4 \cdot 2 + \dots + 4 \cdot \frac{M}{2} + 4 \cdot \frac{M}{2} + \dots + 4 \cdot 2 + 4 \cdot 1 \\ &= 4 \cdot 2 \cdot (1 + 2 + \dots + \frac{M}{2}) \\ &= M \cdot (M + 2) \\ &= O(M^2), \end{aligned}$$

그리고 M 이 홀수일 때는

$$\begin{aligned} & 4 \cdot 1 + 4 \cdot 2 + \dots + 4 \cdot \frac{M}{2} + \dots + 4 \cdot 2 + 4 \cdot 1 \\ &= 4 \cdot (2 \cdot (1 + 2 + \dots + \frac{M-1}{2}) + \frac{M}{2}) \\ &= (M-1) \cdot (M+1) + 2 \cdot M \\ &= O(M^2) \end{aligned}$$

이다. 따라서 격자의 크기가 $M=r \times c$ 인 마이크로어레이 이미지의 격자 주소를 결정하는 알고리즘의 시간 복잡도는 $O(M^2)$ 이다.

물론 입력 데이터인 마이크로어레이 이미지의 상태가 아주 좋지 않은 경우에는 그래프 탐색 알고리즘에 기반한 격자 주소 할당 알고리즘을 적용하기에는 무리가 있다. 이런 경우에는 마이크로어레이 이미지의 부영역의 크기와 반점사이의 평균 거리를 이용하여 일률적인 산술 연산에 의한 주소 할당을 하여 주어야 하는데 이런 방식을 "solid gridding"이라고 하자. 그러면 우리가 주어진 마이크로어레이 이미지에 대해서 어떤 경우에 그래프 탐색 알고리즘에 기반한 방식을 사용하고 어떤 경우에는 "solid gridding"을 할 것인지를 결정해야 하는데 우리는 이것을 이미지의 격자 정규도를 측정하여 해결하고자 한다.

4. 격자 구조 이미지의 정규도

이전에 기술한 바와 같이 격자 구조를 가진 이미지 데이터가 주어지면 이 이미지 데이터의 정규도(regularity)

를 측정하여 "solid gridding"으로 주소 할당이 가능한지 또는 그래프 탐색 알고리즘에 기반한 주소 할당을 사용해야 하는지의 여부를 결정하는 방법에 관하여 설명한다. 격자 구조를 가진 이미지 데이터를 통해서 유전자 발현 정도를 측정하기 위해서는 먼저 각 유전자들이 붙어 있는 위치(반점의 위치)를 정해야 한다. 마이크로어레이 이미지 데이터 같은 경우에서 붙여놓은 유전자 위치는 일반적으로 상당한 오차를 가지고 있어서 사전에 마이크로어레이 장비의 핀(pin)이 찍히는 위치를 알고 있다고 하더라도 스캔되어진 이미지상의 위치와는 많이 차이가 난다.

완전 자동화된 분석 시스템들도 기본적으로 격자 주소 결정을 하기 위한 기준이 되는 반점들이 필요로 하는데 일반적으로는 강한 발현 반응을 보이는 반점(strong target spot)들을 이용한다. 본 논문에서는 강한 발현 반응을 보이는 반점을 정점으로 하는 ϵ -그래프라는 구조를 제시하여 자동으로 반점들의 격자 주소를 찾아주는 알고리즘에 관하여 설명하였다. 격자 구조를 가진 모든 이미지 데이터들에 대해서 이와 같은 방법을 적용시키려고 하는데 실제 실험 결과,

- 강 발현 유전자의 개수가 아주 작은 경우,
- 강 발현 유전자의 개수가 많아도 전체 격자 구조에서 각 격자 주소간의 단위 거리를 파악할 수 있는 영역이 존재하지 않는 경우에는 제한한 방법으로는 격자 주소 결정 과정이 불가능하였다. 이와 같은 경우에는 이미지의 가로-세로로 부영역의 해상도(가로-세로 반점의 개수)에 따라 일정 간격의 격자를 놓고 각 반점들의 격자 주소를 가장 가까운 격자 칸(cell)에 맞추어 주는 "solid gridding" 방법을 사용할 수 밖에 없다. 그러면 어떤 경우에 이전에 제시한 그래프 탐색 알고리즘에 기반한 주소 결정 방법을 사용하고 어떤 경우에는 "solid gridding" 방법을 사용할 것인가 생각해 보자. 본 논문에서는 이러한 문제에 대해서 마이크로어레이의 이미지 데이터 각각에 대해서 정규도(regularity)를 측정하여 어느 값(δ) 이상인 경우에는 그래프 탐색 알고리즘에 기반한 주소 결정 방법을 사용하고 그렇지 않은 경우에는 "solid gridding"을 적용하려고 한다. 이미지 데이터상의 반점들의 정규도라는 것은 반점들의 위치가 "solid grid"와 유사한 정도를 말한다.

격자 구조를 가진 주어진 이미지 데이터 I 의 격자 해상도가 $n \times m$ 이라고 하면 $N = \max\{n, m\}$ 이라 하면 $N = (n-1) \cdot m + n \cdot (m-1)$ 이다. 그리고 c_{ij} 는 정점 u_{ij} 의 상하좌우 방향의 간선들, $u_{ij}^u, u_{ij}^d, u_{ij}^l, u_{ij}^r$, 중에서 다음

을 만족하는 간선들이 횡축 또는 종축에 대해서 기울어진 기울기의 합을 나타낸다고 하자.

$$d_\mu - r_\mu < |e_{ij}^{(u,d,l,r)}| < d_\mu + r_\mu$$

여기서 d_μ 와 r_μ 는 반점 사이의 평균거리와 반점의 평균 반지름을 각각 나타내고, $e_{ij}^{(u,d,l,r)}$ 은 정점 u_{ij} 의 이웃한 4개의 정점간의 간선을 의미한다.

I 의 격자 구조를 가장 잘 표현하고 있는 $G_\epsilon = (V, E)$ 은 간선의 개수가 많고 각각의 간선들이 종축 또는 횡축과 각각 평행한 경우이다. 따라서 I 의 정규도 $r(I)$ 은 다음의 식으로 계산 가능하다.

$$r(I) = \begin{cases} 0 & \text{if } |E| = 0 \\ \sum_{i,j} \frac{a - c_{ij}}{N} & \text{otherwise} \end{cases}$$

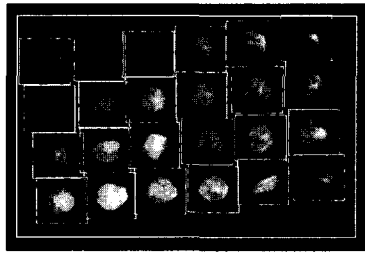


그림 6 기울어진 이미지에 대한 처리시 발생하는 문제점 : 반점들의 분포가 기울어져 있으나 이미지 상에서 격자 주소에 할당된 영역은 x 축과 y 축에 대해서 각각 평행인 상태를 유지하고 있다.

여기서 a 는 격자 구조의 기울기의 최대 값을 의미하며 실험에서는 10° 를 사용하여 계산을 하였는데 마이크로어레이 핀이 유전자 샘플을 놓는 위치의 오차율이 아주 크다고 하여도 일반적으로 10° 를 넘지 않기 때문이다.

이미지의 어떤 영역을 격자 주소로 할당할 때 일반적으로 횡축과 종축에 각 변이 각각 평행인 사각형 모양을 사용하고 있는데 그림 6에서처럼 이미지의 반점들의 분포가 기울어져 있는 경우에는 정확한 영역을 구했다고 할 수 없을 것이다. 이런 경우에 반점들의 분포에서 기울어진 정도를 측정하여 이미지를 회전하지 않고 구하면 더 정확한 결과를 얻을 수 있을 것이다. 이는 이미지를 회전하는 경우에는 이미지 픽셀은 정수형인데 회전은 실수형 계산이므로 발생할 수 있는 오차를 없애는 장점을 가진다. 이에 대한 처리를 위해서 Steinfath 등은 격자의 영역을 표시하는 사각형 상의 4변의 기울기를 평균한 값을 전체 격자의 기울기로 계산하는 방법을 제시하였다[6]. 이런 방법은 마이크로어레이의 위치

오차율(positional error rate)이 낮은 경우에는 간단하게 해결할 수 있는 방법이나 위치 오차율이 높은 경우에 4개의 꼭지점의 위치가 어긋나는 경우에는 실제 기울기에 대한 오차도 크게 나타나는 단점을 가지고 있다. 본 논문에서는 이런 단점을 극복하기 위해 격자 구조내의 각 정점들을 연결하고 있는 간선들의 평균 기울기를 사용하는 새로운 방법을 제시한다.

격자 구조의 기울어진 정도를 측정하는 방식은 본 논문에서 제시한 $G_\epsilon = (V, E)$ 을 이용하는데 $G_\epsilon = (V, E)$ 는 ϵ 의 값에 따라 변하는 그래프이므로 격자 구조를 가장 잘 나타내고 있는 ϵ 값을 구하고 이때의 격자 구조의 기울기를 측정한다. 격자 구조를 잘 표현하고 있는 격자도(degree of grid), $R(G_\epsilon)$ 는 다음과 같이 측정한다:

$$R(G_\epsilon) = \frac{4 \cdot m \cdot n}{\sum_{i=0}^m \sum_{dir=1}^4 q(v_i, dir)}$$

여기서 $|V|$ 는 $G_\epsilon = (V, E)$ 의 정점의 개수를 의미하고 $q(v_i, dir)$ 는 v_i 와 이웃한 4개의 정점간의 격자 주소 값의 차이를 나타내는데 $G_\epsilon = (V, E)$ 상에서 간선으로 연결되어 있는 경우에는 상하좌우의 어느 한 방향으로의 연결성을 의미하는데 완벽한 격자 구조를 나타내는 $G_\epsilon = (V, E)$ 에서는 어느 한 정점과 그 정점의 이웃한 4개의 정점간의 격자 주소 값의 차이는 1일 것이다. 따라서 $q(v_i, dir)$ 는 다음과 같이 계산한다:

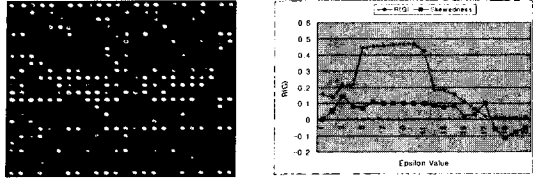
$q(v_i, dir) = c_0 \cdot |h_r(v_i) - h_r(P(v_i, dir))| + c_1 \cdot |h_c(v_i) - h_c(P(v_i, dir))|$ 여기서 $h_r(v_i)$ 는 정점 v_i 의 격자 주소 중 행(row)의 값을 의미하며, $h_c(v_i)$ 는 열(column)의 값을 의미한다. $P(v_i, dir)$ 는 정점 v_i 의 dir 방향의 이웃한 정점을 의미하는 것으로 상은 $dir=1$, 하는 $dir=2$, 좌는 $dir=3$, 그리고 우는 $dir=4$ 이다. c_0 와 c_1 는 두 개의 격자 주소가 같은 방향성을 가지면 가중치가 없이 단순한 절대치이고 다른 방향성을 가지면 가중치를 부여하기 위한 것으로 그 값은 dir 의 값에 따라 변화한다:

$$c_0 = \begin{cases} 1 & \text{if } dir = 1, 2 \\ c & \text{otherwise} \end{cases}$$

$$c_1 = \begin{cases} c & \text{if } dir = 3, 4 \\ 1 & \text{otherwise} \end{cases}$$

여기서 c 는 가중치 값으로 실험에서는 $c = 20$ 을 사용하였다.

그림 7은 NIH(National Institutes of Health)에서의 실험에 의한 마이크로어레이 이미지와 ϵ 의 변화에 따른 격자도의 변화를 보여주고 있는데 격자의 기울기는 격자도가 최대가 되는 ϵ -그래프를 이용해서 측정한다.



(a) NIH 이미지 [12] (b) ϵ 의 변화에 따른 격자도의 변화와 이미지의 기울기를 보여주고 있다. 그림 7 마이크로어레이 이미지의 ϵ 의 변화에 따른 격자도의 변화와 이미지의 기울기를 보여주고 있다: (a) NIH 마이크로어레이 이미지로 Blood B cells에 anti-IgM+CD40L low 48h의 외부 조건을 주고 실험한 이미지의 일부이다. (b) ϵ 의 변화에 따른 격자도의 변화와 $\epsilon = 10$ 에서 격자도가 최대가 되고 이미지의 기울어진 정도는 시계방향으로 0.1° 정도 기울어져 있음을 알 수 있다.

5. 실험 결과

실험에서 사용한 격자 구조를 가진 이미지 데이터는 크게 3가지의 집합으로 구분할 수 있다.

- 1) 가상 데이터(ART0, ART1, ..., ART10) : 가상의 데이터로 완벽한 격자 구조를 가지고 있는 것을 0° 에서 10° 까지 1° 단위로 기울인 데이터이다.
- 2) 마이크로어레이 데이터(NIH1, NIH2, NIH3) : NIH(National Institutes of Health) 데이터 3개로 일반적인 마이크로어레이 데이터이다.
- 3) CCD 데이터(CCD1, CCD2, CCD3) : 그림 8과 같은 플레이트에 베타 조직을 올려 놓고 CCD(Charge Coupled Device) 카메라로 촬영한 이미지 데이터이다. 이들 데이터는 어느 분자 생물학과 연구실에서 생성한 것으로 여러 외부 조건에 의한 베타 성장 실험에 관한 다양한 결과이다.

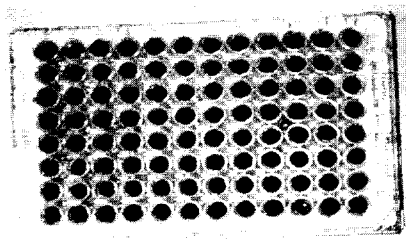


그림 8 96-well 플레이트에 베타 조직을 심어 놓은 모습으로 실제 실험에서는 24-well로 줄여 CCD카메라로 촬영한 데이터를 사용하였다.

그림 9는 이들 17개 중에서 4개의 데이터를 보여주고 있다.

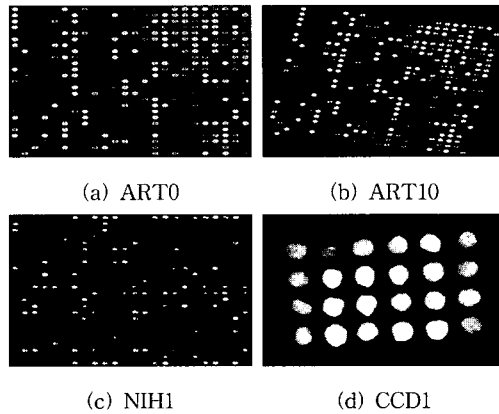


그림 9 실험에서 사용한 격자 구조를 가진 이미지 데이터: (a) 가상의 데이터로 완벽한 격자 구조를 가지고 있고 기울어짐이 없다. (b) 가상의 데이터로 시계 방향으로 10° 의 기울어짐을 가지고 있다. (c) NIH 마이크로어레이 이미지 데이터로 일반적인 두개의 채널(red-green channel) 이미지이다. (d) CCD 데이터로 24-well 플레이트상의 베타 조직 이미지 데이터이다.

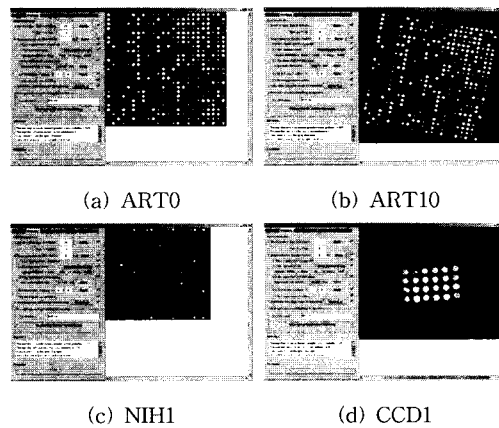


그림 10 4개의 이미지 데이터에 대한 격자 주소 결정을 하고 난 이후의 모습

본 논문에서 제시한 가중치 있는 너비 우선 탐색 기법을 적용하여 격자 주소를 결정하는 방식에 의한 결과는 그림 10과 같다. 그림 10에서 4개의 데이터에 대한 실험에서 모든 격자 주소를 결정한 모습을 각 반점의

영역을 표시한 것으로 확인할 수 있다.

그리고 그림 11은 가상의 데이터에 대해서 정규도를 측정 한 결과인데 그림 11 (a)는 이전에 설명한 가상의 데이터이고 그림 11 (c)는 정규도가 낮게 나오는 새로운 가상의 데이터이다. ϵ 값의 변화에 따른 정규도의 변화를 확인해 볼 수 있는데 그림 11 (a)는 정규도가 ϵ 의 값이 3이상인 경우에는 격자 주소 결정이 가능할 정도의 값을 보이나 그림 11 (c)는 어떤 ϵ 값을 사용한다고 할 지라도 격자 주소 결정이 불가능함을 알 수 있다. 따라서 이런 경우에는 본 논문에서 제시한 그래프 탐색 알고리즘에 의한 격자 주소 결정은 불가능하고 "solid gridding" 방식을 사용해야 함을 알 수 있다.

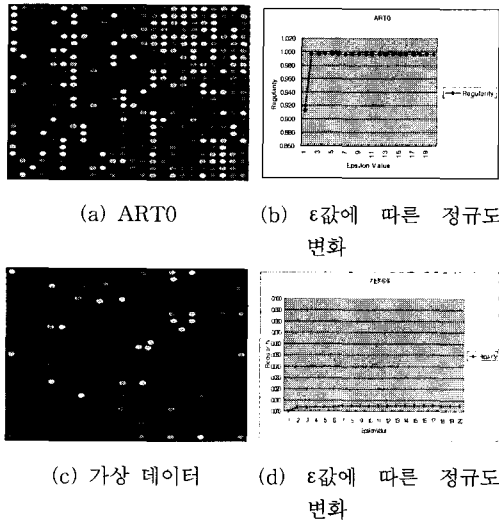


그림 11 2개의 가상의 이미지 데이터에 대한 정규도 측정 결과

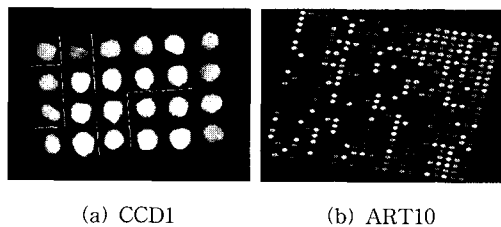


그림 12 기울어진 이미지에 대해서 격자의 정확한 위치와 모양을 계산한 결과 : (a)의 CCD1 이미지는 시계 반대 방향으로 3.6° 기울어져 있고 (b)의 ART10 이미지는 시계 방향으로 10° 기울어져 있다.

마지막 실험은 격자 구조가 기울어진 이미지 데이터에 대한 실험을 하였는데 이번 실험은 CCD 카메라로 촬영한 이미지 데이터 3개와 가상의 데이터를 1°에서 10°까지 기울인 데이터와 NIH 데이터 하나에 대해서 실험을 하였고 표 1은 이들 데이터에 대한 기울어진 정도에 대한 결과를 그림 12는 기울어진 데이터에 대해서 정확한 위치와 모양의 격자를 계산한 결과를 보여주고 있다.

표 1 이미지 데이터에 대한 기울어진 정도를 측정 한 결과: +는 시계방향 그리고 -는 시계 반대 방향으로 기울어져 있음을 의미한다.

데이터	행	열	# 정점	# 간선	격자도	ϵ value	기울기
ART1	25	24	599	1149	0.96	2	1.009
ART2	25	24	559	1149	0.96	3	2.048
ART3	25	24	559	1149	0.96	3	2.048
ART4	25	24	559	1149	0.96	3	4.003
ART5	25	24	559	1149	0.96	4	5.015
ART6	25	24	559	1149	0.96	5	6.013
ART7	25	24	559	1149	0.96	5	7.012
ART8	25	24	559	1149	0.96	7	7.98
ART9	25	24	559	1149	0.96	7	8.719
ART10	25	24	559	1149	0.96	7	9.323
NIH 1	24	24	449	850	0.926	9	0.076
CCD1	4	6	24	38	0.49	4	-3.6
CCD2	4	6	25	39	0.392	9	-5.335
CCD3	4	6	24	38	0.41	7	-5.725

6. 결론 및 향후 연구 과제

본 논문에서는 마이크로어레이 이미지와 같은 격자 구조를 가지고 있는 이미지 데이터를 분석하기 위하여 기존의 이미지 처리 기법이 아니라 이미지 데이터를 ϵ -그래프라는 격자 구조에 적합한 그래프 구조로 변환한 후 그래프 탐색 알고리즘을 이용하여 이미지 상의 반점의 해당 격자 주소를 결정하는 새로운 알고리즘을 제시 하였다. 본 논문에서 제시하는 방법을 사용하여 가상의 데이터와 실제 마이크로어레이 실험을 통해서 생성된 이미지, 그리고 CCD 카메라로 촬영한 격자 구조를 가진 이미지 데이터에 대해서 각각 실험을 하였다. 그리고 제시한 알고리즘의 적용을 위한 이미지의 정규도 (regularity) 분석과 기울어진 이미지 처리를 위해 이미지 자체를 기울이지 않고 또 기존의 이미지 처리 방법에 의한, 예를 들면 프로파일링(profileing) 기법을 사용하지 않

고 ϵ -그래프를 이용한 기울기 측정으로 이미지의 기울기를 계산한 후, 격자 구조를 기울여서 해당 영역의 화소들의 강도를 측정하는 새로운 방법도 제시하였다.

본 논문에서 제시한 격자 주소 결정 알고리즘의 특징은 다음과 같다.

- 1) 최소한의 사용자 개입으로 격자 구조의 이미지 데이터 분석을 위한 격자 주소 결정 과정을 수행한다.
- 2) 마이크로어레이 이미지뿐만 아니라 다른 격자 구조를 가진 모든 이미지 데이터 분석에 그대로 적용 가능하다.
- 3) 격자 구조 이미지 데이터의 정규도(regularity)에 대한 측정으로 자동 처리 가능한 것과 그렇지 않은 것에 대한 구분이 가능하다.
- 4) 격자 구조의 기울기를 기존의 이미지 처리 기법, 예를 들면 프로파일링(profileing)과 같은 방법을 사용하지 않고 구할 수 있다.
- 5) 기울어져 있는 이미지에 대해 실제 이미지를 기울이지 않음으로 해서 이미지 회전 에 따른 오차를 제거할 수 있다.

그리고 향후 연구과제는 다음과 같다.

- 1) 격자 주소 결정에 대한 품질 제어 방법 : 격자 주소와 해당 이미지 영역에 대한 품질 제어에 관한 내용이다.
- 2) 격자 영역에 대한 화소별 강도 측정에 대한 품질 제어 방법 : 화소별 강도 측정 시 배경 강도를 제거하게 되는데 배경 강도가 지역별로 편차가 있으므로 이에 대한 품질 측정 방법에 관한 내용이다.
- 3) 격자 영역에 대한 기하적 분석 : 일반적인 방법은 경계선 탐색과 같은 방법으로 경계선 내부에 대한 화소의 강도를 측정하고 있으나 해당 반점의 기하적 모양을 알 수 없어 다른 이미지 데이터, 예를 들면 단백질 2차원 젤 이미지와 같은 경우에는 기하적 모양도 중요한 특성을 나타내고 있으나 현재의 방법으로는 적용이 불가능하다.
- 4) 정규도가 낮은 이미지 데이터에 대한 격자 주소 결정 방법 : 기존의 "solid gridding" 방법을 개선한 좀더 안정적인 방법론이 필요로 하다.

참 고 문 헌

[1] D. J. Duggan, M. Bittner, Y. Chen, P. Meltzer, and J. M. Trent, "Expression profiling using cDNA microarrays," *Nature Genetics*, Vol.21, pp. 10-14, 1999.

[2] R. Ekins and F. Chu, "Microarrays: their origins

and applications," *Trends in Biotechnology*, Vol.17, pp. 217-218, 1999.

[3] N. Brändle, H. Chen, H. Bischof, and H. Lapp, "Robust parametric and semi-parametric spot fitting for spot array image," *Proceedings of Intelligent Systems for Molecular Biology*, pp. 43-49, 2000.

[4] J. Buhler, T. Ideker, and D. Haynor, "Dapple: Improved Techniques for Finding Spots on DNA Microarrays," *University of Washinton CSE Technical Report UWTR 2000-08-05*, 2000.

[5] C.S. Brown, P. C. Goodwin, and P. K. Sorger, "Image metrics in the statistical analysis of DNA microarray data," *Proceedings of National Academy of Sciences of the United States of America*, Vol.98, pp. 8944-8949, 2001.

[6] M. Steinfath, W. Wruck, H. Seidel, H. Lehrach, U. Radelof, and J. O'Brien, "Automated image analysis for array hybridization experiments," *Bioinformatics*, Vol.17, pp. 634-641, 2001.

[7] <http://rana.stanford.edu/software/>.

[8] M. Eisen, P. Spellman, P. Brown, and D. Botstein, "Cluster analysis and display of genome-wide expression patterns," *Proceedings of National Academy of Sciences of the United States of America*, Vol.95, pp. 14863-14868, 1998.

[9] <http://www.biodiscovery.com/>.

[10] <http://www.mediacy.com/arraypro.htm>.

[11] M. Schena, *Microarray Biochip Technology*, Eaton Publishing, 2000.

[12] <http://lmp.nih.gov/lymphoma/data/rawdata/>.

부 록 : 표기 정의

표 기	정 의
I	마이크로어레이 칩 이미지 데이터
$I_{i,j}$	마이크로어레이 이미지 I 의 (i, j) 위치의 픽셀 값
$G_\epsilon = (V, E)$	하나의 정점이 상하좌우의 방향으로 최대 4개의 간선을 가지는 그래프로서 간선의 연결성은 x 축 또는 y 축 방향으로 ϵ 범위 내에 존재하는 정점들 중에서 가장 가까운 정점을 연결한 것이다.
v^u, v^d, v^l, v^r	$G_\epsilon = (V, E)$ 에서 어느 정점 v 의 이웃한 상, 하, 좌, 우 4개의 정점들
$u_{ij}^u, u_{ij}^d, u_{ij}^l, u_{ij}^r$	$G_\epsilon = (V, E)$ 에서 어느 정점 u_{ij} 의 상, 하, 좌, 우 방향의 간선들
$ e(u, v) $	$G_\epsilon = (V, E)$ 에서 특정한 간선 $e(u, v)$ 의 u 와 v 사이의 거리
$\{u_{i,j}\}$	마이크로어레이 이미지 I 상에 존재하는 모든 정점의 집합
$r(I)$	마이크로어레이 이미지 I 에 있는 정점 $\{u_{i,j}\}$ 의 정규도
c_{ij}	정점 u_{ij} 의 상하좌우 방향의 간선들의 기울기의 합
$R(G_\epsilon)$	$G_\epsilon = (V, E)$ 의 격자도(degree of gridding)



정 호 열

1997년 부산대학교 전자계산학과 졸업(학사). 1999년 부산대학교 전자계산학과 졸업(이학석사). 1999년 ~ 현재 부산대학교 전자계산학과 박사과정. 관심분야는 생물정보학, 유전자 클러스터링, 그래프 이론, 이미지 처리



황 미 녕

2000년 부산대학교 전자계산학과 졸업(학사). 2000년 ~ 현재 부산대학교 전자계산학과 석사 과정. 관심분야는 생물정보학, 유전자 클러스터링



유 영 중

1996년 부산대학교 전자계산학과 졸업(학사). 1998년 부산대학교 전자계산학과 졸업(이학석사). 2002년 부산대학교 전자계산학과 졸업(이학박사). 2002년 ~ 현재 부산외국어대학교 컴퓨터공학과 전임강사. 관심분야는 컴퓨터 그래픽스



조 환 규

1984년 서울대학교 계산통계학과 졸업(학사). 1986년 한국과학기술원 전자계산학과 졸업(공학석사). 1990년 한국과학기술원 전자계산학과 졸업(공학박사). 1990년 ~ 현재 부산대학교 전기전자정보컴퓨터공학부 교수. 관심분야는 그래프 이론, 생물정보학, 그래픽스