

## 3D 형광이미지 분석을 위한 레인 검출 및 추적 알고리즘

이복주 · 문 혁 · 최영규<sup>†</sup>

<sup>†</sup>한국기술교육대학교 대학원 컴퓨터공학부

### Lane Detection and Tracking Algorithm for 3D Fluorescence Image Analysis

Bok Ju Lee, Moon Hyuck and Young Kyu Choi<sup>†</sup>

<sup>†</sup>Korea University of Technology and Education, School of Computer Science and Engineering

#### ABSTRACT

A new lane detection algorithm is proposed for the analysis of DNA fingerprints from a polymerase chain reaction (PCR) gel electrophoresis image. Although several research results have been previously reported, it is still challenging to extract lanes precisely from images having abrupt background brightness difference and bent lanes. We propose an edge based algorithm for calculating the average lane width and lane cycle. Our method adopts sub-pixel algorithm for extracting rising-edges and falling edges precisely and estimates the lane width and cycle by using k-means clustering algorithm. To handle the curved lanes, we partition the gel image into small portions, and track the lane centers in each partitioned image. 32 gel images including 534 lanes are used to evaluate the performance of our method. Experimental results show that our method is robust to images having background difference and bent lanes without any preprocessing.

**Key Words** : lane detection, lane tracking, 3D Fluorescence image, PCR gel electrophoresis image, k-means algorithm, projection profile, sub-pixel algorithm

#### 1. 서 론

PCR (Polymerase Chain Reaction : 중합효소 연쇄 반응) 젤 전기영동(Electrophoresis)은 크기에 기초하여 생체분자들을 분리하고, 효율적으로 정량적인 정보를 추출하는 가장 일반적인 방법 중의 하나이다. 생체분자는 젤 전기영동 실험 동안에 특정효소의 작용에 의해 다수의 DNA 단편들로 분해될 수 있다. 폴리 아크릴아미드(polyacrylamid) 또는 아가로스 젤(agarose gel)에 전류를 흘려주면, DNA 단편들이 각각 다른 속도로 분산되고, 시간이 지나면 젤이 염색된다[1]. 이와 같은 전기 영동 기술은 탄소나노튜브의 성장과 같은 다양한 분야에서도 활용되고 있다[2].

Fig. 1은 전기영동 완료된 아가로스 젤을 촬영한 영상의 예를 보여준다. 검사 대상 조각 분자들은 전류의 흐름 방향을 따라 수직 하단으로 이동하고, 검사 대상

물질의 구성원들을 나타내는 작은 사각형(밴드)들이 수직 일직선으로 나열한 형태를 가진다. Fig. 1에서 보이는 밴드들의 수직 나열 모음을 레인(lane)이라 부르고 검사대상 물질들을 구별하는 기준이 된다. 그리고 레인 내에서 보이는 작은 사각형 모양의 수평 밴드(band)들의 위치는 검사 대상 물질을 구성하는 분자들

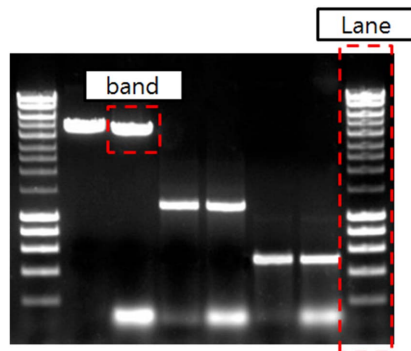


Fig. 1. An Example of electrophoresis gel images.

<sup>†</sup>E-mail : ykchoi@koreatech.ac.kr

의 이동거리 값을 표현하고, 수평 밴드의 넓이는 이동된 구성 분자들의 축적량을 나타낸다.

염색된 젤은 디지털 영상으로 기록된다. 디지털 이미징 기법의 출현은 분석과 후속 프레젠테이션(subsequent presentation)에서 젤 영상을 획득하는 직접적인 방법을 제공하고 있다.

디지털화 된 젤 전기영동 영상은 많은 분자 생물학 응용프로그램에서 생물학적 물질로부터 다양한 정보를 추출하는데 사용된다. 예를 들어, 알려지지 않은 샘플과 알려진 샘플의 관계는 영상에서 DNA의 패턴을 비교함으로써 달성된다[3]. 그러나 사람의 육안을 통한 영상의 수동적인 평가는 대량의 반복적인 전기영동 젤 분석 작업을 위해 많은 시간이 소요된다. 또한, 평가자의 개인경험에 의존하기 때문에 단조롭고, 주관적이고, 오류를 발생하기 쉽다. 이에 따라, 부정확한 육안 검사의 문제점을 해결하기 위한 다양한 자동화 방법들이 연구되고 있다.

본 논문의 구성은 다음과 같다. 먼저, 2장에서 관련된 연구를 살펴보고, 3장에서 새로운 레인 검출 알고리즘을 제안한다. 4장에서 휘어진 레인 추적 방법을 소개하고, 5장에서는 실험결과를 보이며, 6장에서 결론을 맺는다.

## 2. 관련 연구

자동화의 가장 중요한 단계는 레인을 정확하게 식별하는 것이다. 레인의 식별을 위해서 영상에서 레인을 정확히 검출하는 레인 검출(lane detection)이 반드시 필요하다[4]. 또한, 휘어져 있는 레인들의 정확한 검출을 위해 레인의 중심을 추적하는 레인추적(lane tracking) 단계도 필요하다.

Fig. 1에서 볼 수 있듯이 젤 영상에 있는 레인들은 일반적으로 폭과 간격이 일정하고 밴드와 배경간의 밝기에 어느 정도 차이가 있는 특성을 가지고 있다. 이 특성을 이용한 프로젝션 프로파일 기반의 방법[5-8]들이 연구되고 있다. Ye 등은 원형의 구축 요소에 의한 모폴로지의 폐쇄(closing)를 적용하여 배경을 산정하고 이를 원 영상에서 빼는 'Top-hat Transform'을 수행하여 점진적으로 변화하는 배경을 제거한 후, 프로파일의 에지를 분석하여 레인을 추출한다[9]. 그러나 이 방법은 구축요소의 모양과 크기에 따라 결과 영상에서의 밝기 변화가 크고, 밴드의 밝기에 왜곡이 생길 수 있다는 단점이 있다. Machado 등은 영상을 최소 필터로 전처리한 후, 수직으로 프로젝션한 결과를 최소나 중간 값으로 필터링하여 잡음을 제거한 다음, 1차 모멘트를 이용

한 반복적인 기법으로 레인의 중심축을 추출한다[1]. 그러나 이 방법은 레인 사이의 배경 구간을 분리하지 못하고 레인의 중심축만을 추출하며, 다양한 영상처리에 의해 밴드 밝기에 왜곡이 생길 가능성이 크다는 문제점을 가진다. Akbari 등은 반자동화된 레인 추적 방법을 제안했는데, 이 방법은 추적결과를 편집해야만 한다[10]. Park 등은 휘어져 있는 레인을 검출하기 위한 방법을 제안하였는데, 전체 영상의 수직 프로젝션 프로파일 계산을 통해 평균 레인 폭을 추정하고, 이 값을 이용하여 레인을 추적한다[11]. 그러나 이 방법은 배경영역이 밝기차이를 가지는 영상에서는 정확한 결과를 갖기 어렵고, 학습 단계에서 결정을 위한 대부분의 파라미터 세트의 크기에 상대적으로 둔감하다는 단점이 있다.

본 논문에서는 PCR 전기영동 영상에서 자동으로 레인을 안정적으로 검출하고 중심을 추적하는 방법을 제안한다. 특히, 에지 기반의 알고리즘을 통해 배경영역의 밝기차이가 있어도 안정적으로 중심을 추출할 수 있으며, 부화소 알고리즘을 사용하여 보다 정밀한 계산이 가능하며 휘어진 레인에 대해서도 잘 대응할 수 있도록 하였다.

## 3. 제안된 레인 검출 알고리즘

Fig. 2는 제안된 레인 검출 및 추적 알고리즘을 보여주고 있다. 레인 검출을 위해서는 입력 영상에서 평균

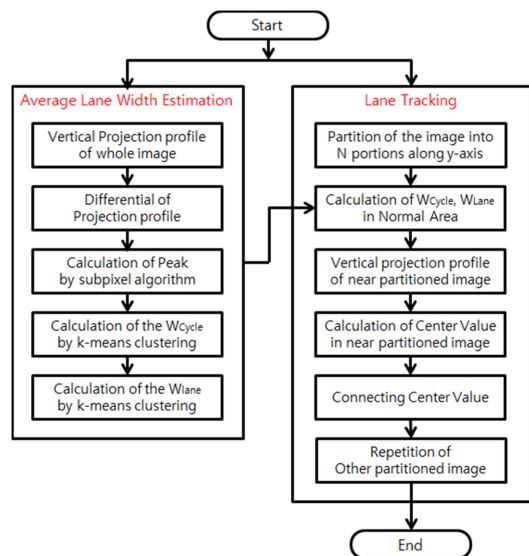


Fig. 2. Flowchart of the proposed scheme.

레인 폭( $W_{Lane}$ )과 주기( $W_{Cycle}$ )를 정확히 검출해야 한다. 이 장에서는 젤 영상에서 안정적으로 레인 폭과 주기를 검출하는 방법을 소개한다.

**3.1. 수직 프로젝션 및 미분**

Fig. 1을 보면 젤 영상의 특성을 이용하여 레인 검출에 수직 프로젝션이 사용될 수 있음을 알 수 있다. 영상의 수직 프로젝션은 각 열의 화소 값을 누적하여, 2차원 정보를 1차원으로 변환한다. Fig. 3(a)는 이 영상의 수직 프로젝션 프로파일을 보여주는데, 값들이 급격하게 증가하거나 감소하는 부분들이 반복적으로 나타난다. 이것은 영상에서 밝기 변화가 큰 에지(edge) 영역에 해당하는데, 이들이 레인의 시작과 끝 정보를 나타내고, 이들 사이의 거리를 이용하여 레인 폭과 주기를 추정할 수 있다.

프로젝션 프로파일에서 에지의 위치를 추출하기 위해 프로젝션 프로파일의 미분을 구한다. 미분에는 다양한 필터를 사용할 수 있는데, 제안된 방법에서는 Fig. 4와 같은 필터를 사용하였다. Fig. 3(b)는 미분 결과를 보여주고 있다.

**3.2. 에지 위치의 정밀 검출**

미분 프로파일에서 만들어지면 에지 위치를 검출해야 한다. 에지는 미분 프로파일에서 국부 최대치와 최소치(local maximum, minimum)로 나타나는데 이를 피크(peak)라 하자. Fig. 3을 보면 국부 최대치는 레인의 시작 에지이고, 국부 최소치는 레인이 끝나는 부분이다.

이와 같은 미분 프로파일의 피크 위치는 정수값으로

나타난다. 따라서 보다 정밀도를 향상시키기 위해 보간법(interpolation)을 사용하였다. 이것은 부화소(sub-pixel) 처리라고 하는데, 화소 단위의 정밀도에 비해 5배 이상 정밀한 것으로 알려져 있다. Fig. 5는 이러한 부화소 처리의 개념을 보여주고 있다.  $E(i)$ 가 최대값이라고 할 때, 실제의 최대값은  $i$ 가 아니라  $i$ 와  $i+1$  사이에 존재하는데 샘플링에 의해  $i$  위치가 사라진 것으로 생각할 수 있다. 따라서 다음과 같이 더 정밀한 최대값의 위치를 계산할 수 있다.

$$i_{max} = i - \frac{E(i+1) - E(i-1)}{4E(i) - 2E(i+1) - 2E(i-1)} \quad (1)$$

**3.3. 평균 레인 폭과 주기 계산**

에지 위치가 정밀하게 검출되면 이들을 이용해 평균 레인 폭과 주기를 계산해야 한다. 검출된 에지 위치는 다양한 잡음을 포함하고 있다. Fig. 6(a)와 같이 누락 레인을 포함하는 영상이나, Fig. 6(b)와 같이 잡음에 의해 에지가 발생할 수 있다. 따라서 이들에게서 불필요한 에지를 제거하고 안정적으로 레인의 폭과 주기를 계산해야 한다. 이를 위해, 우리는 k-means 알고리즘을 사용한다.

k-means 알고리즘은 데이터의 군집화(clustering)을 위해 사용되는데, 다음과 같은 평균자승오차(Mean Squared Error) 함수  $I_{MSE}$ 를 반복적으로 적용하여 전체 오차를 최소화시키는 방법이다.

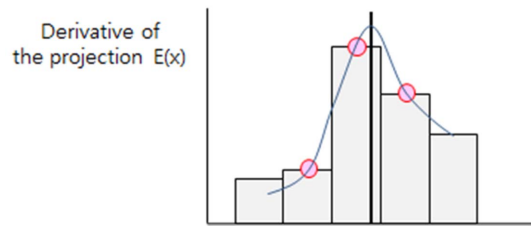


Fig. 5. The concept of interpolation.

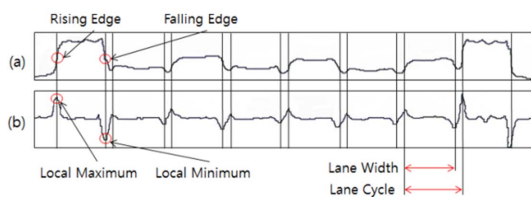


Fig. 3. (a) Projection profile, (b) Projection differential.

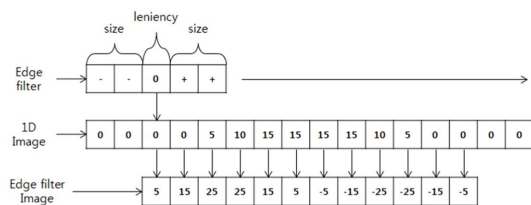


Fig. 4. Configuration of the Edge filter.

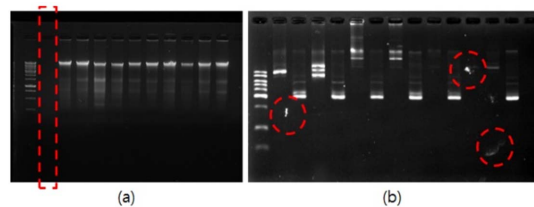


Fig. 6. (a) An example of Image containing the noise, (b) An example of Image containing the missing lane.

$$J_{MSE} = \sum_{i=1}^k \sum_{x \in \omega_i} |x - \mu_i|^2 \quad (2)$$

$$\text{where } \mu_i = \frac{1}{N} \sum_{x \in \omega_i} x$$

k-means 알고리즘은 먼저 데이터 집합  $\{x_1, \dots, x_n\}$ 으로부터 임의의 K개의 벡터를 선택하여 k개의 초기 중심집합  $\{y_1, \dots, y_n\}$ 을 만든다. 만약 데이터  $x_n$ 이  $y_i$ 에 가장 가깝다면 클러스터  $x_i$ 에 속하도록 한다. 결국, 데이터 집합이 k개의 클러스터들  $\{x_1, \dots, x_k\}$ 로 나누어진다.

$$x_i = \{x_n | d(x_n, y_1) \leq d(x_n, y_i), i = 1, \dots, k\} \quad (3)$$

군집화 단계에서 구한 새로운 그룹들에서 각각의 중심을 갱신한다.

$$y_i = c(X_i) = \frac{1}{M} \sum_{m=1}^M \{X_i\}, i = 1, \dots, k \quad (4)$$

데이터와 가장 가까운 클러스터 중심들과 거리의 합으로 총 왜곡(distortion)을 구한다.

$$D = \sum_{n=1}^N d(x_n, y_{i(n)}) \quad \text{where } i(n) = k, \quad (5)$$

$$\text{if } x_n \in X_k$$

총 왜곡이 변하지 않거나, 설정된 반복횟수에 도달할 때까지 앞의 단계들을 반복한다.

k-means 알고리즘 적용하기 위해 먼저 상승 에지와 하강 에지의 쌍들을 모두 찾아 각각의 길이를 구한다. 이 데이터를 k-means 알고리즘에 적용하고 가장 큰 그룹의 평균값을 평균 라인 폭으로 간주한다. 이렇게 되면 잡음에 의한 에지들의 영향이 없어지고 정상적인 라인들에 의해 폭이 구해진다.

마찬가지로, 상승 에지와 인접하는 다른 상승 에지의 쌍을 구하고, 그 거리를 모두 찾아 k-means 알고리즘에 적용한다. 역시 가장 큰 그룹의 평균 거리가 라인의 평균 주기가 된다.

라인의 폭과 주기를 계산하면 영상에서 각각의 라인들을 구분할 수 있게 되며, 각 라인의 중심을 구할 수 있다.

## 4. 라인 추적 알고리즘

가끔 젤 영상에서 라인이 휘어진 경우가 발생한다. Fig. 7(b)는 라인이 휘어져 있는 젤 영상을 보여주는데, 이 경우 앞에서 구한 폭과 주기만을 이용하여 모든 영역의 라인을 정확히 분리기는 어렵다. 따라서 이러한 영상에서는 앞에서 구한 라인 정보를 바탕으로 라인의 중심을 추적할 필요가 있으며 이를 라인 추적(lane tracking)이라 한다.

라인 추적은 먼저 영상을 수직으로 여러 영역으로 분할하는 것으로 시작한다. 각 부분 영상에서 구한 폭과 주기가 전체 영상에서의 결과와 잘 맞는 영역들을 선택하고 이들을 중심으로 인접한 분할 영역의 라인 위치를 갱신하는 방법을 사용한다

### 4.1. 영역 분할 및 정상영역 설정

먼저 영상을 수직으로 N개의 영역으로 분할한다. 분할 영역은 수평으로 길쭉한 영상이 된다. 실험에서는 N을 8로 하였다. 분할 영상들 중에서 중심이 휘어져 있지 않은 정상 영역을 찾아야 한다. 정상 영역은 앞에서 추출한 폭과 주기에 잘 맞는 부분이 된다. 분할 영역들 중에서 정상 영역들을 선택하기 위해 앞 장에서 설명한 방법을 분할 영상에 다시 적용하고 폭과 주기를 계산한다. 이 결과가 전체 영상에서 추출한 결과와 유사하면 정상 영역으로 선택된다. 물론, 사용자가 정상 영역을 지정하는 것도 가능하다. Fig. 8의 영상에서는 (b), (c), (d)가 정상 영역으로 선택되었다.

### 4.2. 인접 분할 영역의 라인 중심값 추출

정상 영역이 선택되면 이제 이웃한 부분 영역에서 라인의 중심값을 갱신한다. 이를 위해 부분 영역에 대한 프로젝션 프로파일을 구하고, 미분 프로파일을 구한 후 피크를 찾는다. 이때, 상승 피크와 하강 피크를 구분해야 한다. 다음으로 인접한 정상 영역에서의 상승-하강 피크 쌍과 비교하여 라인의 이동 정도와 라인 중

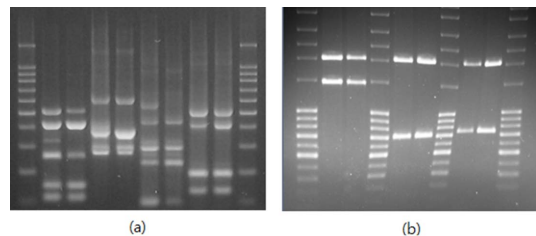


Fig. 7. (a) An example of Normal image, (b) An example of image containing the curved lane.



심을 계산한다. 이동시 레인의 폭이 변경될 수 있으므로 변경된 폭도 함께 저장하고, 이 과정을 반복하여 영상의 상단과 하단의 모든 부분 영상들을 처리한다.

이때, Fig. 8에서 볼 수 있듯이 분할영역 안에 흰 밴드 영역이 있을 수도 있고 없을 수도 있다. 밴드가 존재하는 경우는 문제가 없지만, 없는 경우에는 상승-하강 피크 쌍이 나타나지 않는다. 이 경우는 가장 인접한 처리된 영역에서 구한 레인의 정보를 유지하도록 한다.

### 5. 실험결과 및 분석

제안된 알고리즘은 Windows-7 운영체제 하의 Pentium-PC에서 구현되고 실행되었는데, C++를 이용하여 구현되었으며 Visual Studio 2010 환경에서 Release 모드로 컴파일 되었다. Fig. 9는 실험을 위해서

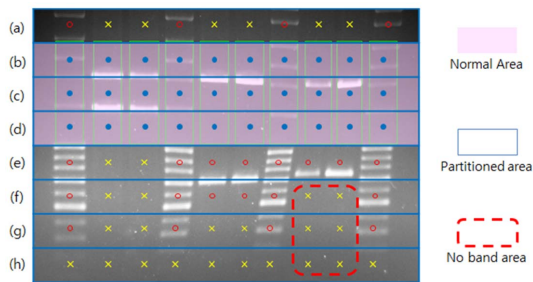


Fig. 8. Image the user sets the normal area.

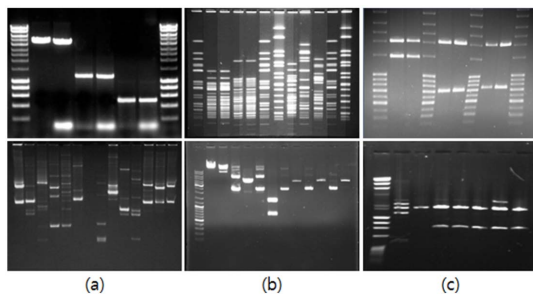


Fig. 9. Examples of Test Image, (a) Normal image, (b) Background Brightness Difference, (c) Curved Lnae.

사용된 영상들을 보여주고 있다. 실험에서는 534개의 레인을 가지는 32개의 PCR Gel 영상을 사용하였다. 일부 젤 영상은 Fig. 9(b)와 같이 배경영역의 밝기가 부분적으로 변하는 경우가 있다. 이런 영상을 처리하기 위해서 배경제거 기법을 사용하는 전처리를 사용할 수도 있지만, 제안된 방법에서는 에지 기반 알고리즘을 통해 전처리 없이도 안정적으로 레인을 검출한다. 또한, Fig. 9(c)와 같이 일부 젤 영상은 휘어져 레인을 포함하기도 한다. 전체 실험 영상들에는 배경영역이 밝기차이를 가지는 10개의 PCR Gel 영상과 휘어져 있는 레인을 가지는 6개의 PCR Gel 영상이 포함되었다.

Fig. 10 은 배경영역이 밝기차이를 가지는 PCR Gel 영상에 대한 레인 자동 검출결과를 보여주고 있다. Fig. 10의 영상은 모두 배경영역이 밝기차이를 가지고 있는 영상이지만, 제안된 방법은 이러한 영상에 대해서도 안정적으로 레인을 검출하는 것을 알 수 있다. 또한, 사용자의 간섭 없이도 자동적으로 레인을 검출한다.

Fig. 11 은 휘어져 있는 레인을 포함하는 PCR Gel 영상에 대한 레인 추적 결과를 보여주고 있다. 휘어져 있는 레인을 포함하는 영상은 레인의 구부러짐 현상이 영상의 어느 부분에서 나타날지 예측할 수 없기 때문에 사용자로 하여금 정상영역을 선택하게 한다. Fig. 11의 오른쪽의 초록색 영역만큼 정상영역을 설정하게 되면 분할된 영역의 레인중심을 찾아가며, 레인을 안정적으로 추적하는 것을 확인할 수 있다

Table 1은 실험에 사용된 영상들의 레인 검출 및 레인 추적 성공 비율을 보여주고 있다. 16개의 Normal 영상에 대해서는 97%의 확률로 레인 검출을 성공하였으며, 10개의 배경영역의 밝기차이가 존재하는 영상에 대해서도 94%의 확률로 안정적으로 레인을 검출하는 것으로 나타나고 있다. 또한, 휘어져 있는 레인을 포함

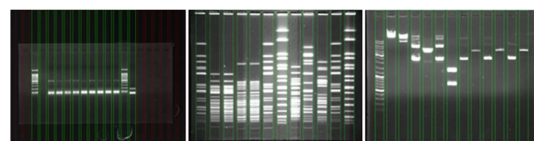
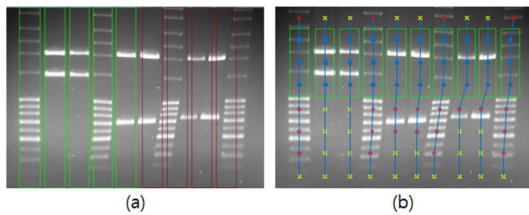


Fig. 10. Result by proposed scheme in image that having Brightness difference.

Table 1. Performance of image that having Brightness difference

	Number of Lanes	Number of Detected Lanes	Rate
Normal	276	268	97.0%
Background Brightness Difference	177	167	94.3%
Curved Lanes	81	73	90. 1%



**Fig. 11.** Result by proposed scheme in image that having Curved lanes.

하는 6개 영상의 레인 추적에 대해서도 90%의 확률로 레인을 검출하는 것으로 나타났다. 입력 영상의 레인 개수가 너무 적거나, 레인의 밴드영역이 너무 적은 경우에는 상승-하강 에지 계산을 위한 데이터가 충분하지 않기 때문에 레인검출 및 추적에 오류가 발생하였다.

이와 같은 실험을 통해 제안된 방법이 PGR Gel 영상에서 안정적으로 레인을 검출하고, 추적하는 매우 효과적인 기법임을 확인할 수 있었다.

## 6. 결론 및 향후 과제

본 논문에서는 3D 형광이미지 분석을 위해서 PCR 젤 전기영동 영상으로부터 DNA 지문을 분석하기 위한 새로운 레인 검출 및 추적 알고리즘을 제안하였다. 제안된 방법은 영상의 수직 프로젝션 프로파일에서 피크를 정밀하게 찾고 상승 에지와 하강 에지를 추출한다. 이들은 k-means 알고리즘을 통해 군집화되는데, 이 그룹들 중에 최대 그룹을 찾고 이를 이용해 평균 레인 폭과 레인 주기를 계산한다. 제안된 방법은 에지 기반 알고리즘으로 이는 배경영역의 밝기차이가 비교적 큰 영상에 대해서도 매우 안정적으로 레인을 검출할 수 있었다. 본 논문에서는 휘어져 있는 레인에 대해 추적 방법도 제시하였는데, 영역 분할을 통해 레인 중심값을 안정적으로 추적할 수 있었다.

향후 현재 실험에 사용된 검정색 배경의 흰색의 밴드를 가지는 영상뿐만 아니라 다양한 색상의 영상들에 대한 실험이 진행될 예정이다. 배경의 밝기 차이가 지나치게 크거나, 심하게 흐린 영상에 대한 전처리 방법에 대한 연구도 추가되어야 한다.

## 참고문헌

1. Machado, Alexei, et al. "An iterative algorithm for segmenting lanes in gel electrophoresis images." *Computer Graphics and Image Processing. Proceedings., X Brazilian Symposium on. IEEE, (1997).*
2. Chang, Han-Beet, et al. "Field emission properties of carbon nanotubes grown on micro-tip substrates using an electrophoretic deposition method" *Journal of the Semiconductor & Display Technology, 9(4), pp. 7-12, (2010)*
3. Hoelzel, A. Rus, and Gabriel A. Dover. "Molecular genetic ecology." *IRL Press at Oxford University Press, (1991).*
4. Wong, Richard TF, et al. "LaneRuler: automated lane tracking for dna electrophoresis gel images." *IEEE transactions on automation science and engineering, 7.3, pp. 706-708, (2010).*
5. Kim, Taek Hyeon, et al. "Automatic DNA Image Recognition System for Diagnosis of Tuberculosis." *Journal of Korea Multimedia society, 2009.5, pp. 722-725, (2009).*
6. Kim Seung Il, et al. "Automatic Alignment of Electrophoresis Gel Images Based on Standard Deviation of Vertical Profiles." *Journal of KISS : Software and Applications, 39.8, pp. 631-638, (2012).*
7. Lee, Wan Yeon, et al. "Automatic Analysis Scheme for Multiple Images of Ongoing Electrophoresis Gel." *Journal of KISS : Software and Applications, 39.8, pp. 672-677, (2012)*
8. Lee, Jiann-Der, et al. "Automatic DNA sequencing for electrophoresis gels using image processing algorithms." *Journal of Biomedical Science and Engineering, 4.08, pp. 523, (2011).*
9. Ye, Xiangyun, et al. "A recent development in image analysis of electrophoresis gels." *Vision Interface'99, Trois-Rivières, 19.21, pp. 432-438, (1999).*
10. Akbari, et al. "Automatic lane detection and separation in one dimensional gel images using continuous wavelet transform." *Analytical Methods, 2.9, pp. 1360-1371, (2010).*
11. Park, Sang Cheol, et al. "Lane detection and tracking in PCR gel electrophoresis images." *Computers and electronics in agriculture, 83, pp. 85-91, (2012).*

접수일: 2016년 2월 17일, 심사일: 2016년 3월 3일,  
게재확정일: 2016년 3월 23일