

# 3차원 연결 성분 레이블링 알고리즘에 관한 연구

최익환<sup>한국인제대학교</sup>, 이병일<sup>한국인제대학교</sup>, 최현주<sup>한국인제대학교</sup>, 최홍국<sup>한국인제대학교</sup>  
인제대학교 컴퓨터공학부

## A study on 3D connected component labeling algorigm

IkHwan Choi, ByoungIl Lee, HyunJu Choi, HeungKook Choi  
School of Computer Engineering, Inje Univ.

### 요 약

볼륨데이터에서 관심대상의 특징을 추출하기 위해서 3D레이블링을 3차원 세포영상의 분석에 적합한 레이블링 방법인 SIL(slice Information base labeling)을 제안하였다. SIL은 각 슬라이스 정 보를 이용하여 레이블링을 수행하므로 영상의 특징에 맞는 레이블링으로의 확장이 유용하고 메 모리 효율이 높다. 몇 개의 실험 영상으로 다른 방법과 비교한 결과 성능면에서도 우수한 결과를 얻었다. 또한 레이블링을 통해서 얻어진 피쳐값으로 세포 영상을 분석하였으며, 콘포컬 현미경 영상을 이용하였을때 실험영상에서 결과를 추출하는데 걸린 시간은 SIL방법이 기존 방법보다 2배 가량 빨랐다. 다양한 3차원 레이블링 방법 중 적용되는 영상에 따라 각기 다른 결과를 얻었지만, 3차원 세포영상의 분석에는 SIL 방법이 우수하다는 결론을 얻었다.

### 1. 서론

레이블링을 통해 오브젝트를 분리하는 것은 영상분석에서 중요한 과정 중의 하나인데, 3차원 데이터로 구성되는 많은 데이터에서 3차원적인 오브젝트의 영상분할이 요구되고 있다. 라벨은 서로 떨어진 component들을 구별해 내는 방법을 말하며 물체분석, 영상분할 및 병합, 잡음제거, 인식 등 여러 분야에 활용되고 있다[1,2]. 최근 의료기기의 발달로 2차원 영상뿐만 아니라 CT, MRI나 LSCM같은 장비에서 일어지는 3차원 볼륨데이터에서 오브젝트의 분리에 대한 연구가 요구되고 있다. 일반적으로 3차원데이터의 처리시간은 2차원영상보다 상당히 크므로, 볼륨데이터에서 빠르고 효율적인 레이블링 알고리즘이 필요하다.

볼륨데이터의 오브젝트에 대해 영상분할을 한 후 레이블링하는 방법은 2차원 단면 영상에서 오브젝트를 분리해 내기 위해 4방향이나 8방향의 마스크를 이용하는 방법과, 6방향, 18방향, 그리고 26 방향의 3차원 마스크를 사용하여 3차원데이터를 한번에 접근하여 처리하는 방법이 있는데, 3차원 마스크를 사용하는 방법은 처리시간이 길어지고 사용하는 구조체에 따라 메모리를 많이 필요로 하는 단점이 있다. Jiuxiang Hu는 DL(dynamic list)를 사용하여 3차원 빠른 레이블링을 시도하였다[6].

본 연구에서는 기존의 레이블링 방법에서 개선한 레이블링 방법을 몇 가지 영상에 대해 실험하고 그 결과를 비교해 보았다. 볼륨데이터들은 2차원 영상들을 쌓아놓은 구조라 할 수 있다. 이를 이용하여 2차원 데이터의 레이블링 된 슬라이스로부터 3차원 레이블링을 처리하는 방법을 제안한다.

## 2. 관련연구

NFL(Non-recursive flood-filling labeling)에서는 현재 위치 및 종료 위치를 설정하고 방문 flag를 설정할 수 있는 LS(label stack)구조체를 영상의 크기 만큼 선언한 다음 현재 위치에 대하여 현재 라벨번호를 불인다[5].

```
class LS
{
    bool isVisited;
    position pReturnPos;
}
```

그 다음 주변 픽셀 중에서 방문하지 않은 1인 픽셀 n을 찾는다. 이때 n이 존재 할 경우 픽셀n의 복귀위치를 현재 픽셀 x의 위치로 설정해둔다. 그 다음 현재 픽셀의 위치를 다시 n의 위치로 설정하여 이를 반복함으로써 flood-filling이 구현된다.

NFL알고리즘을 수정하면 쉽게 3D 레이블링으로 확장 할 수 있는데 스캔 방향에 z축을 추가하고 연결 정보를 4/8방향에서 6/18/26방향으로 확장하고 현재 위치정보나 복귀위치가 복셀정보를 가지도록 수정하면 볼륨데이터에서의 레이블링이 가능해진다.

Jiuxiang Hu가 제안한 GCCL(Greedy Connected Component labeling)은 한번의 스캔으로 레이블링을 수행하며 DL를 사용한다[6]. 따라서 NFL처럼 LS구조체를 선언할 필요가 없어 메모리 활용면에서 뛰어나다. GCCL은 이진화된 영상을 스캔하면서 라벨이 없는 복셀을 찾으면 6/18/26 방향으로 연결된 모든 복셀이 레이블링이 될 때까지 라벨을 붙여나간다. 이렇게 라벨 붙여진 복셀들은 각 라벨마다 각각의 DL에 속하는 component가 된다. 그러나 복셀이 중복 사용되거나 라벨이 재할당 되는 것을 막기 위하여 4-value flag를 사용한다.  $t(i,j,k)$ 를 component 복셀이라 하면, 0,1,2,3들의 값을 가질 수 있다. 만약  $t(i,j,k)=0$  이라면 배경에 해당하는 값이고 이 복셀의 값은 변하지 않는다.  $t(i,j,k)=1$  이면 목표 픽셀에 해당하는 값으로 라벨값이 할당 되지 않았고 새로운 DL 시작이 될 수 있음을 의미한다.  $t(i,j,k)=2$  이면 DL에 추가되는 새로운 노드를 의미한다. 하지만 새로운 라벨이 될 수

없다.  $t(i,j,k)=3$  이면 이미 DL에 추가되었으며 새로운 노드가 될 수 없음을 의미한다.

## 3. 제안하는 방법

3차원 데이터의 레이블링 알고리즘을 위해 먼저 2D 레이블링 방법으로 CCCL(contour based connected component labeling)을 적용하여 각 슬라이스에서 오브젝트 정보들을 추출하는 방법을 제안한다. 이 CCCL은 [4]에서 제안한 레이블링을 응용하는 방법으로 정수형 영역 정보를 갖는 영상을 오직 한번 스캐닝함으로써 모든 영역들의 정보를 추출할 수 있다.

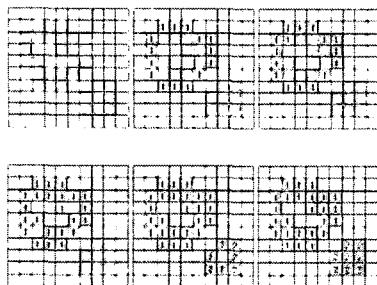
제안된 방법은 오브젝트의 영역의 경계를 추적하면서 가장 왼쪽의 픽셀에만 라벨을 붙여주고 그 픽셀의 라벨값으로 오브젝트의 라벨을 점차적으로 매겨나간다. 하지만 가장 왼쪽의 픽셀 값만 라벨을 붙여주고 나가는 알고리즘은 추후 라벨이 붙을 것을 감안하면 중복연산이 되고 watershed등의 전처리를 거친 영상이 아닌 경우 외곽선을 끝까지 추적해야 되므로 알고리즘이 복잡해지므로 본 연구에서 이 단계의 방법을 외곽선 추적을 통한 라벨할당으로 전처리 등에 의존하지 않게 하였다. 개선한 CCCL 레이블링 방법은 다음과 같다.

입력 영상과 동일한 크기의 버퍼를 할당하고 그 값을 모두 0으로 초기화한다. 각 영역들은 1부터 시작하여 차례대로 라벨이 증가하며 매겨지게 된다. 만약 현재 스캔하고 있는 위치의 라벨 값이 0이면 경우 3가지로 나누어 처리된다. 먼저 바로 왼쪽에 있는 점의 픽셀 값과 현재 픽셀 값이 동일한 경우, 바로 왼쪽의 픽셀과 동일한 라벨을 매긴다.

두 번째 현재 좌표에서 좌측 상단에 존재하는 픽셀의 값과 현재 픽셀의 값이 같지 않은 경우, 라벨 값을 1 증가시키고 현재 영역의 경계를 추적하면서 동일한 라벨을 적용한다.

세 번째, 위의 두 경우를 만족하지 않는 경우는 훌(hole) 추적을 시작한다. 이 때, 새로운 오브젝트로 인식하지 않도록 주의의 라벨을 검사하여 동일한 라벨을 매긴다. 스캔시 만약 현재 스캔하고 있는 위치에 이미 라벨이 매겨져 있다면 다음 픽셀로 이동한

다[그림 1].

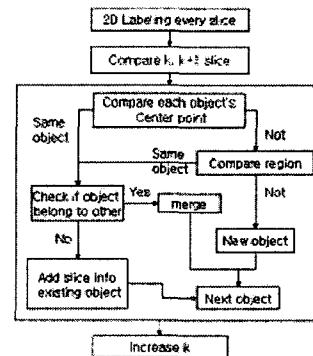


[그림 1] CCCL 진행과정

볼륨데이터에서의 레이블링을 위해 본 논문에서 제안하는 새로운 방법인 SIL(Slice Information based Labeling)은 슬라이스들의 레이블링된 정보를 이용하여 3D 레이블링 한다. 볼륨데이터는 슬라이스 데이터의 집합과 같다[3]. 따라서 전체 볼륨을 스캔할 필요 없이 슬라이스에서 레이블링된 정보를 바탕으로 연관관계를 찾아줌으로써 3D 레이블링을 할 수 있다. 또한 SIL에서는 볼륨데이터를 읽어 들이고 각 슬라이스 별로 2D 레이블링을 수행하므로 영상에 맞는 레이블링 알고리즘을 선택하여 사용할 수 있다는 장점이 있다. 제안한 SIL 알고리즘은 다음과 같다.

K번째 슬라이스의 object정보들과 k+1번째 슬라이스 object정보들을 비교하고 같은 3차원 object에 해당하는지 검사하여 볼륨데이터의 레이블링을 구축한다. 이때 필요한 레이블링 정보로는 오브젝트의 중심점, 영역크기, id가 필요하다. 비교하는 방법은 먼저 해당하는 k, k+1번째 object들 영역의 교집합 여부를 판단한다. 교집합이 생기지 않을 경우 비교대상이 아니다. 겹쳐질 경우 중심점에 해당하는 픽셀들이 연결되었는가를 비교한다. 만약 도넛 모양처럼 중심점에 해당하는 픽셀이 비어있거나 중심점에 해당하는 픽셀이 연결되지 않았을 경우에는 각 object들이 픽셀들을 조사하여 연결 점을 형성하는가를 비교한다. Object들이 연결 점을 가지지 않을 경우에는 각 object들은 다른 object들이며 k+1번째 object의 경우 새로운 label이 된다. 이러한 방법은 중심점을 먼저 확인하기 때문에 대부분의 경우 연산량이 O(1)이다. 다음 그림 2은 SIL방법에 대한 알고리즘 흐름도이다.

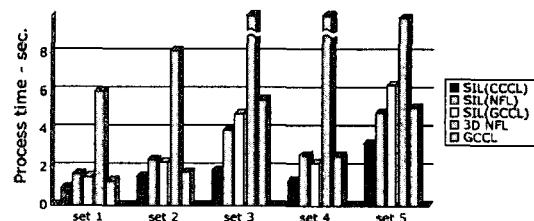
이 알고리즘은 k번째와 k+1번째를 비교함으로 이루어지는데 이전에 새로운 라벨로 선언되었지만 그 이후에 같은 개체임이 밝혀지면 라벨 충돌이 발생할 수 있다. 주로 혈관영상 같이 나뭇가지 형태를 가진 개체일 경우에 많이 발생한다. 이 경우 merge를 수행하여 기존 라벨에 재적용한다. 이 merge과정은 3D 라벨 정보를 DL로 저장하는 과정이다. 각 3D 개체는 각각의 DL에 추가되게 된다. 즉 다른 DL로 저장되었던 개체를 하나의 DL로 만들어주는 과정은 하나의 DL에 다른 DL 포인터를 연결 시켜줌으로써 쉽게 해결 할 수 있다.



[그림 2] SIL 알고리즘

#### 4. 결과

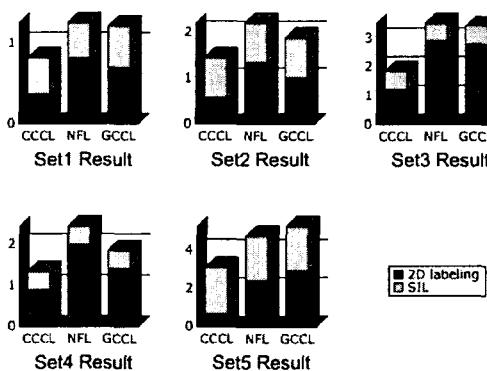
3D 레이블링에서의 결과비교는 SIL방법에 여러 레이블링 알고리즘: CCCL, NFL, GCCL을 사용한 방법들과 3D로 확장한 3DNFL, GCCL방법, 5가지를 컨포렐 현미경에서 얻어진 영상 5set으로 각각 비교하였다. 각 방법들의 수행 시간은 알고리즘의 수행과 연결된 개체들의 정보를 얻고, 일련의 메모리 할당 등의 연산처리가 포함된 총 처리시간이다.



[그림 3] 다섯가지 방법의 영상 적용 결과 비교

그림 3에서 결과를 보면 set1의 경우 SIL(CCCL) 0이 0.937초, SIL(NFL) 1.657초, SIL(GCCL) 1.516초, 3D NFL 5.984초, GCCL 1.297초의 시간을 보이며, 제안한 SIL방법이 전체적으로 빠름을 알 수 있다. 3D NFL의 경우 알고리즘 자체만을 보면 빠른 시간을 보여주나, LS구조체 등의 선언으로 동적 메모리 할당을 사용함으로 인해 특히 할당을 위한 프로세스 타임이 소모가 크며, 메모리 할당량도 커서 효율적인 면에서 좋지 못하기 때문에 이를 위한 개선이 요구된다. GCCL의 경우 전반적으로 좋은 성능을 보이지만 메모리 할당부분에 있어서 개선점이 이루어진다면 좀 더 빠른 시간을 보일 수 있다.

링에 대해 다른 접근방법으로 Slice-based 3D 레이블링 알고리즘을 제안하였으며, 다른 방법들에 비해 좋은 결과를 얻을 수 있었다. 3D NFL의 메모리로 인해 전체적으로 저조한 성능을 보였고 GCCL의 경우 전반적으로 좋은 성능을 보이지만 메모리 할당부분에 있어서 개선점이 이루어진다면 좀 더 빠른 시간을 보일 수 있다. 반면 SIL의 경우 한 장의 슬라이스 크기만큼만 메모리 할당이 이루어지고 재사용 되므로 메모리 소비가 작고, 메모리 할당으로 인한 프로세스 타임이 없어진다. 전반적으로 CCCL을 이용한 SIL방법이 좋은 성능을 보일 뿐만 아니라 시스템적으로도 효율적인 알고리즘이라 할 수 있다.



[그림 4] 각 영상의 결과비교

그림 4은 SIL에서 각각의 2D 레이블링과 SIL처리 루틴이 걸리는 시간을 측정한 것이다. 즉 그림 10의 set1은 0.937초가 걸리는데 이는 set1 result에서 보듯이 2D 레이블링인 CCCL을 처리하는데 0.36초 SIL 처리 루틴으로 0.45초가 걸렸음을 말한다. 여기에서 남아있는 0.127초는 메모리 할당 등의 기타 연산이 걸리는 시간이다. SIL의 내부 처리 시간을 살펴보면 Set5 Result와 같이 영상이 복잡해 질수록 SIL의 시간 소모가 더 커지는 것을 알 수 있다.

## 5. 결론

본 연구에서는 기존의 블룸데이터에서의 레이블

## [참고문헌]

- A. Rosenfeld, J. Pfaltz, "Sequential operations in digital picture processing", J. ACM 13 (4) (1966) 471-494. October, 1966.
- R. Haralick L. Shapiro, "Computer and Robot Vision Volume", Academic Press, Vol.2, pp.241-242, 1982.
- RAZDAN A. ,PATEL K. , FARIN G. AND CAPCO D. G. 2001. Visualization of Multicolor LCM data set. Computers and Graphics, 25, 3, 371-382.
- Y. Ishiyama, C. Funaoka, F. Kubo, H. Takahashi, F. Tomita, "Labeling Board Based On Boundary Tracking", 11th ICPR, pp.34-38, 1992.
- 김도현, 강동구, 차의영, "비재귀 Flood-Fill 알고리즘을 이용한 적응적 이미지 레이블링 알고리즘", 한국정보처리학회 논문지 제 9-B권 제 3호, pp.337-342, 2002년 6월
- J. Hu, G. Farin, M. H. II, S. P. Massia, G. Nielson, A. Razdan, "Statistical 3D Segmentation With Greedy Connected Component Labelling Refinement", <http://prism.asu.edu/>