

비지도 학습 기반의 TEM 이미지 결손 원자 탐지 알고리즘

김민성, 이상철*

한국기초과학지원연구원

kms00737@kbsi.re.kr, *sclee02@kbsi.re.kr

요약

물성 분석에서 투과전자현미경 이미지를 통한 결손 원자 검출은 전문가의 수작업 판독에 의존하는 한계가 있다. 이러한 점을 해결하기 위해 본 연구에서는 주기적인 원자 배열로 구성된 결정 구조 내 결손 원자를 자동으로 탐지하는 비지도 학습 기반 알고리즘을 제안한다. 제안된 기법은 원본 영상에서 원자 중심점을 추출한 영상을 입력으로 사용한다. 이후 들로네 삼각분할과 필터링 과정을 통해 구조적 연결 정보를 확보하고, 원자 간 최근접 이웃 선분의 개수에 기반한 지역 벡터 패턴을 추출한다. 이렇게 얻어진 패턴들을 일반화하여 대표적인 벡터 집합으로 단순화한 뒤, 각 원자점 주변의 이웃 관계와 대조함으로써 결손 가능성이 높은 위치를 효과적으로 추정한다. 제안된 기법의 성능을 검증하기 위해 TEMImageNet 데이터셋의 다양한 격자 구조를 대상으로 실험을 수행하였다. 그 결과, 결손 비율이 증가하더라도 정밀도가 안정적으로 유지되어 결손 위치 판별의 신뢰성을 확인할 수 있었다.

I. 서론

결정 재료 내 원자 단위의 결함은 물성에 중대한 영향을 미치므로, 이를 탐지하는 기술은 재료 특성 분석에 필수적이다. 이 과정에서 투과전자현미경은 원자 수준의 해상도를 바탕으로 결함 구조를 직접 관측할 수 있는 핵심 도구로 활용되고 있다. 그러나 촬영 조건, 대비의 민감도와 같은 요인으로 인해 영상 내 결함 식별이 어렵다. 또한, 결함 탐지 자동화 도구가 일부 개발되었으나 샘플의 구조적 다양성과 사용자의 요구를 충분히 반영할 수 없기 때문에 실제 현장 분석은 여전히 전문가의 경험에 크게 의존하고 있다.

본 연구에서는 결손 원자를 자동으로 식별하기 위해 비지도 학습 기반의 탐지 기법을 제안한다. 제안하는 방법은 사전 학습 데이터나 결정 구조에 대한 사전 정보 없이, 원자 간의 기하학적 배치를 분석하여 반복적으로 나타나는 격자 패턴을 도출한다. 이후 각 원자 위치에서 도출된 패턴과 대조함으로써, 결손 발생 가능성이 높은 지점을 효과적으로 추출한다. 이러한 접근은 전문가의 경험과 수작업에 의존하던 기존 분석 방식을 대체할 수 있으며, 다양한 결정 구조와 복잡한 재료 시스템에도 확장 가능한 자동화 탐지 기반을 제공한다.

II. 본론

본 연구의 전체 과정은 규칙적인 배열 구조를 갖는 원자 좌표를 기반으로, 결손 가능성이 높은 위치를 추출하는 것을 목표로 한다. 이를 위해 세 단계 과정을 거쳐 결손 원자 위치를 도출한다. 입력 이미지는 256×256 해상도에서 검정 배경 위에 각 원자 중심을 1픽셀 흰점으로 표시한 형태이며, 제안된 방법의 세부 내용은 다음 절에서 설명한다.

2.1. 들로네 삼각분할을 이용한 그래프 생성 및 지역 벡터 추정

규칙적인 원자 중심 좌표에서 결손 위치를 탐지하기 위해서는 점간의 구조적 연결 관계를 파악할 필요가 있다. 들로네 삼각분할은 각 점을 삼각형으로 연결되되 외접원 내부에 다른 점이 존재하지 않도록 함으로써, 점 간의 관계를 격자 구조로 변환하여 구조적 연결성을 나타낸다. 그러나 이 방법은 모든 점을 최소 삼각형으로 연결하기 때문에 그림 1의 (a)에서 보이는 것처럼 불필요한 장거리 연결선이나 규칙적 구조 내의 대각선 변이 포함될 수 있다. 이에 따라, 일정한 면적을 유지하는 규칙적 원자 배열을 얻

기 위해 전역 평균 길이보다 짧은 변만을 남기는 방식을 적용하여 불필요한 연결을 제거한다. 이후 그림 1의 (b)와 같이 구축된 그래프를 바탕으로 각 원자에 대해 인접 원자까지의 벡터를 계산하여 거리와 방향 정보를 포함하는 지역 벡터를 추정한다. 지역 벡터는 점 간의 각도와 규칙적 배열 패턴을 반영하고 있어, 추후 결손 원자 후보 위치를 예측하는 핵심 정보로 활용된다.

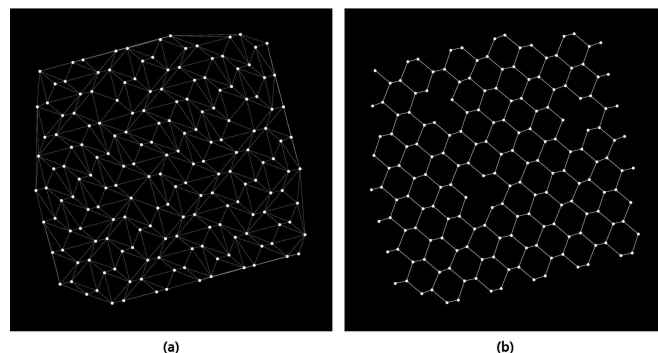


그림 1. (a) 들로네 삼각분할 결과, (b) 평균 선분 길이 기반 필터링 결과

2.2. 들로네 삼각분할 기반 필드 추정을 이용한 저밀도 후보 탐색

모든 원자 점에 대해 지역 벡터를 직접 적용해 결손 후보를 탐색하는 방식은 불필요한 연산이 많아 계산 비용과 시간이 크게 증가할 수 있다. 따라서, 결손 가능성이 높은 영역을 우선 검출해 탐색 대상을 줄이는 과정이 필요하다. 이를 위해 본 연구에서는 들로네 삼각분할 기반 필드 추정(Delaunay Triangulation Field Estimator, DTFE)을 사용한다. DTFE는 들로네 삼각분할을 기반으로 각 삼각형의 면적을 인접 정점에 분배해 국소 밀도를 추정하는 기법으로, 면적간의 밀도 변화를 통해 결손 위치 인근 지점을 효과적으로 탐지할 수 있다. 그림 2의 (a)는 DTFE 결과로, 짙은 회색의 저밀도 영역이 결손 원자 위치와 근접하며, 인접 영역의 결손 후보 점은 그림 2의 (b)에서 X로 표시하였다.

2.3. 결손 위치 추정

앞서 선별한 각 저밀도 후보점에 대해 인접 원자와의 벡터 방향을 분석

하여, 사전에 정의된 지역 벡터 중 가장 유사한 벡터를 선정한다. 이 과정에서 인접 원자와의 방향이 다수 매칭되고 결손 발생 가능성이 높은 벡터를 선택함으로써, 그림 2(b)에 나타난 노란색 삼각형과 같은 결손 후보 위치를 산출한다. 이때, 생성된 후보 위치 중 기존 원자점과 과도하게 가까운 경우는 오탐 가능성이 높으므로 기존 원자 간 최소 평균 거리를 기준으로 필터링을 수행한다. 이후 남은 후보 점들은 위치 중복을 제거하고 대표 좌표를 결정하기 위해 DBSCAN(Density-Based Spatial Clustering of Applications with Noise) 알고리즘을 적용한다. DBSCAN은 인접 밀도 기반의 군집화 방식으로, 노이즈를 배제하면서 인접한 후보 점들을 하나의 군집으로 묶을 수 있어 결손 위치의 안정적인 산출이 가능하다. 이 과정을 통해 최종 후보 점들의 각 군집 중심점을 결손 원자 위치로 확정한다.

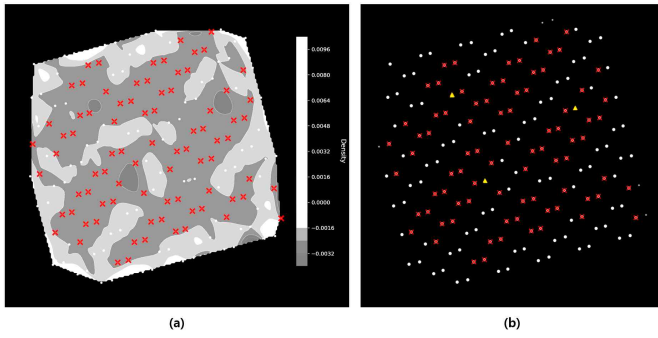


그림 2. (a) DTFE 기반 후보점 생성 결과 (b) 결손 원자 후보 위치 추정 결과 (○: 기존 원자, x: 저밀도 후보 지점, △: 결손 원자 위치 추정 지점)

III. 실험

제안한 결손 원자 탐지 알고리즘의 성능을 검증하기 위해 TEMimagenet [1] 데이터셋을 기반으로 실험을 수행하였다. TEMimagenet에서 규칙적인 원자 배열을 갖는 이미지를 선별하기 위해 원자 개수가 30개 이상인 6,588장을 추출하였다. 성능 평가는 검출된 좌표가 실제 결손 위치로부터 1픽셀 이내에 존재할 경우 정탐(True Positive)으로 판정하는 기준을 적용하였다. 또한, 오탐과 실제 결손 검출 비율을 정량화하기 위해 정밀도(Precision)와 재현율(Recall)을 사용하였으며, 종합적인 평가를 위해 F1-score를 함께 활용하였다.

표 1은 중앙부에 1~3개의 랜덤하게 결손을 삽입한 경우의 결과를 나타낸다. 이 조건에서 전체 합산 기준 정밀도는 0.962, 재현율은 0.952, F1-score는 0.957을 기록하여, 제안된 기법이 규칙적인 원자 배열 구조에서 결손 위치를 높은 정확도와 재현율로 안정적으로 탐지함을 보여준다.

또한, 이미지 내 원자 수 대비 결손 비율을 10%, 20%, 30%로 달리하여 성능을 비교한 결과는 표 2와 같다. 결손 비율이 증가함에 따라 재현율과 F1-score는 점진적으로 감소하는 경향을 보였으나, 정밀도는 모든 경우에서 0.92 이상으로 높게 유지되었다. 이는 제안된 알고리즘이 결손 비율이 높아져도 불필요한 오탐(False Positive)을 최소화하며, 검출한 결손 위치의 신뢰도가 높다는 것을 의미한다. 결손 비율이 높아질수록 재현율이 하락한 이유는, 결손이 많아짐에 따라 격자의 규칙적 패턴이 점점 사라져, 지역 벡터 기반 패턴 분석에서 일부 결손 위치를 탐지하지 못했기 때문이다. 이러한 결과는 규칙적인 패턴이 유지되는 영역에서는 제안된 알고리즘이 항상 높은 예측 신뢰도를 보이며 결손 위치를 정확히 식별할 수 있음을 의미한다.

그림 3은 원본(GT) 이미지와 결손 비율을 10%, 20%, 30%로 변화시켰을 때의 결과를 보여준다. 흰색 동그란 점은 기존 원자 위치를, 노란색 네모 점은 제거된 원자(결손 위치)를 나타낸다.

표 1. 랜덤한 결손 생성 시 성능 결과

vacancy	Precision	Recall	F1-score
1~3	0.962	0.952	0.957

표 2. 원자 수 대비 결손의 비율에 따른 성능 비교

vacancy ratio	Precision	Recall	F1-score
10%	0.955	0.759	0.845
20%	0.948	0.725	0.822
30%	0.922	0.672	0.777

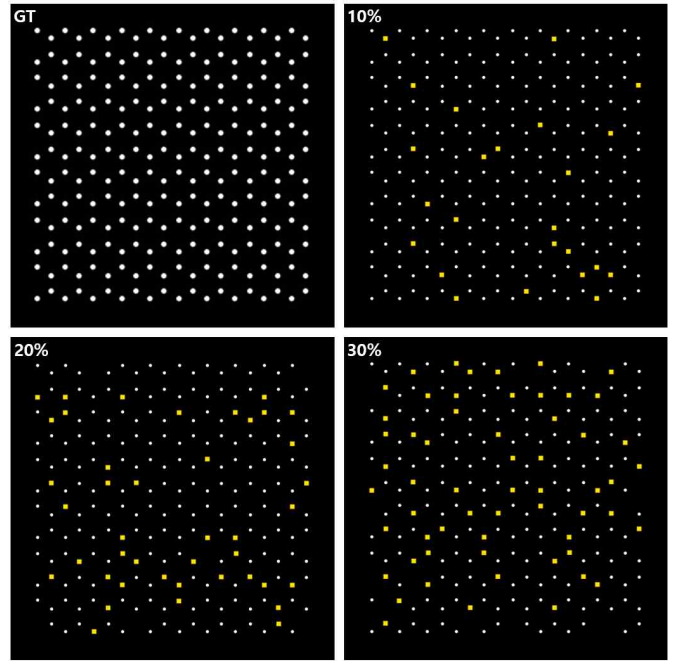


그림 3. 결손 비율에 따른 결손 원자 탐지 결과 (○: 기존 원자, □: 최종 결손 원자 위치 추정 지점)

IV. 결론

본 논문에서는 사전 학습 데이터나 결정 구조에 대한 사전 정보 없이 결손 원자를 자동 탐지할 수 있는 비지도 학습 기반 기법을 제안하였다. 제안 방법은 원자 좌표로부터 지역적인 규칙 격자 패턴을 추출하고, 국소 밀도 분석을 통해 결손 가능성이 높은 영역을 선별한 뒤, 군집화를 통해 최종 결손 원자 위치를 산출하는 방식이다. TEMimagenet 데이터셋을 활용한 실험에서 제안 기법은 높은 정밀도와 재현율을 달성하였다. 이러한 결과는 본 방법이 다양한 규칙적 원자 배열 구조에서 결손 위치를 신뢰성 있게 탐지할 수 있음을 보여준다.

ACKNOWLEDGMENT

This research was supported by the basic science research program (C524200) funded by the Korea Basic Science Institute.

참 고 문 헌

- [1] J. Madsen, M. R. Clausen, T. S. Jespersen, et al., "TEMImageNet: A large-scale multi-domain dataset for electron microscopy image analysis," arXiv preprint arXiv:2003.10585, 2020.