A nanoXRD Based Analysis on HVPE GaN Structure Combined with Machine Learning

Grad. Sch. Eng. Sci., Osaka Univ.[1], JASRI[2]

°(D2) Z. D. Wu[1], (M2) Y. Nakanishi[1], Y. Hayashi[1], T. Tohei[1],

K. Sumitani[2], Y. Imai[2], S. Kimura[2], and A. Sakai[1]

E-mail: sakai@ee.es.osaka-u.ac.jp

[Introduction] GaN has been an attractive material thanks to its exclusive properties, including wide direct bandgap (~3.49 eV), high electron mobility (> 1000 $cm^2/V \cdot s$). With the achievement in high-quality GaN substrate growth, crystal structure evaluation methods are also developed, e.g., multi-photon-excited photoluminescence (MPPL), synchrotron white beam X-ray topography (SWXRT) and nanobeam X-ray diffraction (nanoXRD). However, manual analysis of the experimental data for novelty or defect recognition becomes challenging since the experimental data acquisition rate increases. Here, a novel method based on machine learning is utilized to analyze the clustering properties based on enormous raw nanoXRD patterns, which help us investigate the crystal structural characteristics.

[Experiment] This research focuses on the area shown in Fig. 1 (a) with the size of 30 × 40 μm^2. Figure 1(a) is the cross-sectional MPPL image obtained from an HVPE GaN sample, of which the dark contrast probably represents a relatively lower crystallinity. The nanoXRD measurement was conducted in SPring-8 BL13XU with a beam size of around 700 (hor.)×470 (ver.) nm^2. With the help of Gauss fitting on Intensity-2θ profile, lattice constants were extracted from 2-200 and 2-202 diffractions of total 1271 diffraction patterns taken from the sample with an interval of 1 μm. We applied the unsupervised machine learning algorithm, uniform manifold approximation and projection (UMAP), to learn the clustering properties of the raw nanoXRD patterns and compared the UMAP plots with the lattice constants *c*-*a* plots.

[Results and Discussion] UMAP is an unsupervised dimension reduction algorithm, which reduces the dimensionality of input datasets into a low dimension space with the relative distance between datasets in high-dimensional space remaining. We first study the behavior of UMAP on the raw nanoXRD patterns. Figure 1(b) shows the UMAP plots for the total 1271 patterns distribution in two-dimensional (2D) space. Two clusters are colored dark blue and light blue, compared with the lattice constants *c*-*a* plots and the MPPL image (Figs. 1(c) and (a), respectively). Groups 1 and 2 in the UMAP plots correspond to the groups 1 and 2 of lattice constants plots as shown in Fig. 1(c), respectively. Interestingly, group 1 relates to the lower part of the sample with a rather linear *c*-*a* relationship, which implies a higher crystallinity. Meanwhile, the higher part of the sample marked by group 2 has more discretely distributed *c*-*a* plots, which is probably caused by the defects. The UMAP plots successfully clustered raw nanoXRD patterns with different crystallinity. The presentation section will show more discussions on the mechanism and physical explanation of UMAP plots.

[Acknowledgment] This work was partially supported by JSPS KAKENHI (JP16H06423). We are also grateful for the illuminating discussions with B. Zhang and K. Fukui.

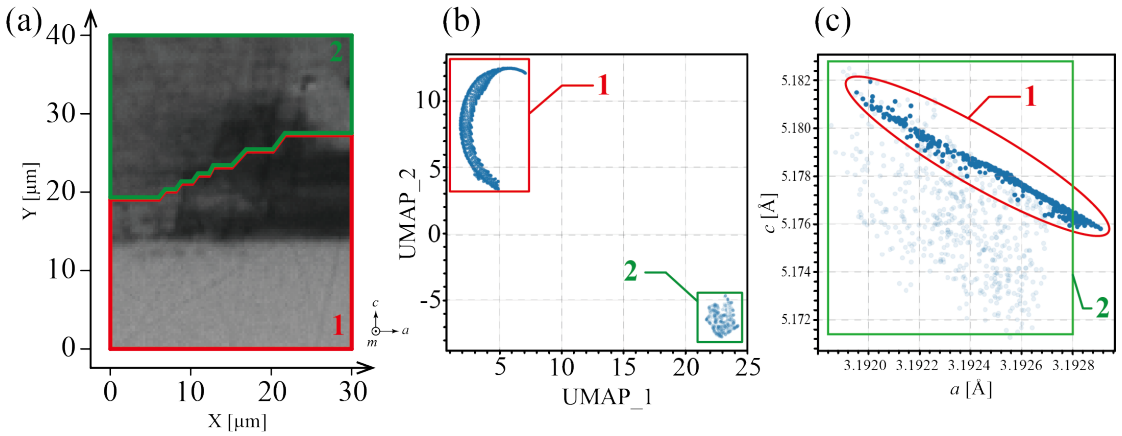

Fig. 1 (a) MPPL image obtained from the cross-sectional *m* plane of the HVPE GaN sample. (b) UMAP plots of total 1271 nanoXRD patterns in 2D space. (c) Lattice constants *c*-*a* plots.