Data-Driven Automated Control Algorithm for Floating-Zone Crystal Growth Using Reinforcement Learning

Harada S[1,2*],Tosa Y[3], Omae R[3], Matsumoto R[3], Sumitani S[3].
*shunta.harada@nagoya-u.jp
1 Institute of Materials and Systems for Sustainability (IMaSS), Nagoya University, Japan
2 Department of Materials Process Engineering, Nagoya University, Japan
3 Anamorphosis Networks, Japan

Materials processes have been upgraded, automated, and efficiently optimized by applying various kinds of informatics algorithms and the field of "process informatics" were emerging. Although the parameters for most materials processes are time-independently given, which means that a set of input parameters gives a single output, some materials processes are manually controlled according to the information obtained during operation. For example, in crystal growth by the optical floating zone (FZ) method, an operator monitors the status of the melt in a furnace by a camera and changes the input parameters to maintain suitable conditions for single-crystal growth. In the present study, we predicted the dynamics of crystal growth by the optical FZ method using a Gaussian mixture model (GMM) [1], and optimized the operation trajectory by reinforcement learning for the automated operation.

There are two characteristics in the operation data of the material manufacturing process, which requires adaptive control according to the process states including FZ crystal growth process. One is that the amount of data for learning is limited, and so-called "big data" cannot be acquired. It is unrealistic to run thousands of experiments for automation. The other is that the operating trajectories are not exactly the same, but they are all similar trajectories. The operation is carried out according to the process states, but it is natural to control with the similar operating trajectories in which the operation was once succeeded. Therefore, it is necessary to construct a prediction model in a limited parameter space with a limited amount of data. Considering these characteristics of the operation data, we selected to apply for the prediction of the dynamics by Gaussian mixture model (GMM) in the present study. GMM is a model that can predict nonlinear dynamics near the training trajectories with high sample efficiency. In order to validate the predication of the dynamics by GMM, we created an emulator program of FZ crystal growth process and generated a virtual operation trajectory. As a result, from only five demonstration trajectories, we successfully predicted the operating dynamics using GMM with better precision than obtained by using linear regression or neural networks.

Once the prediction model for the dynamics of the operation trajectory with high accuracy, it is possible to generate an ideal operation by reinforcement learning. In the present study, the control model was trained using proximal policy optimization (PPO). In reinforcement learning, it is generally important to appropriately set a reward function that quantifies whether the state is good or bad. As a result of the design of the reward function, in order to accurately control the operation of the FZ crystal growth by reinforcement learning from the dynamics predicted by GMM, in addition to the error from the ideal shape of grown crystal considering the necking process, the error from the training trajectory should be included in the reward function. Finally, we succeeded in achieving an operation trajectory that is closer to the ideal shape than when manually controlled by applying GMM and PPO.

References

[1] Omae R, Sumitani S, Tosa Y, Harada S, Prediction of operating dynamics in floating-zone crystal growth using Gaussian mixture model. Sci. Tech. Adv. Mater. Methods 2022;2:294-301.