

Prediction of the Thruster Performance in Hall Thrusters Using Neural Network

IEPC-2017-453

*Presented at the 35th International Electric Propulsion Conference
Georgia Institute of Technology • Atlanta, Georgia • USA
October 8 – 12, 2017*

Hiroataka Fuchigami,¹ Yusuke Egawa,² Taichi Morita³ and Naoji Yamamoto⁴
Kyushu University, Kasuga, Fukuoka, 816-8580, Japan

Abstract: We have been developing a thrust prediction code using neural network for constructing auto controlling system in Hall thrusters. We use 73 training data sets which consists of operational conditions (inner /outer/trim coil current, xenon mass flow rate, discharge voltage) and thrust. The initial value of weights and parameter updating methods affects the reconstruction accuracy and the calculation time. Current code with 5 layers and 100 nodes predicts thrust efficiency with 20 % uncertainty due to overfitting. The number of training data set is only 73, 2 layers and 10 nodes is enough and it can predict re-construct thrust within 5% uncertainty.

Nomenclature

I_{sp}	=	specific impulse
$\dot{m}$	=	mass flow rate of propellant
V_d	=	discharge voltage
I_{in}	=	inner coil current
I_{out}	=	outer coil current
I_{trim}	=	trim coil current
F	=	thrust
SGD	=	stochastic gradient descent
BP	=	back propagation
W	=	weights
ε	=	learning rate
L	=	loss function
$ReLU$	=	rectified linear unit
h	=	sum of squares of gradient
t	=	iteration count
n	=	number of previous nodes

I.Introduction

THRUSTES are essential to keep and change the designed orbit. Recently, the adoption of electric propulsion system for main propulsion system has been increasing in respect of fuel consumption¹⁾. Actually, development

¹ Graduate student, Department of Advanced Energy Engineering Science, fuchigami@aees.kyushu-u.ac.jp.

² Graduate student, Department of Advanced Energy Engineering Science, yusuke@aees.kyushu-u.ac.jp.

³ Assistant Professor, Department of Advanced Energy Engineering Science, morita@aees.kyushu-u.ac.jp

⁴ Professor, Department of Advanced Energy Engineering Science, yamamoto@aees.kyushu-u.ac.jp

of the electric propulsion with high efficiency and large thrust for all-electric satellite is proceeding at rapid pace in Japan ²⁾.

The asteroid explorer “Hayabusa” demonstrated superiority of electric propulsion; its fuel consumption is 5 to 10 times better than that of chemical propulsion ³⁾. Therefore, it’s possible to drastically reduce the fuel that accounts for a high percentage of the gross weight of planetary probes and satellites. One disadvantage of electric propulsion is that thrust is small due to limitation of electric power available in outer space. For that reason, the thrust to power ratio, which is the thrust per unit electric power, is an important indicator in electric propulsion.

Hall thrusters are promising thruster, since they are well balanced in specific impulse (Isp) and thrust to power ratio compared with other electric propulsion ^{4, 5)}. Hall thruster also can be expected to achieve life time, and the thrust efficiency can be 50% or more at specific impulse of 1500 to 2000 seconds. As example of installation in the spacecraft, the main propulsion system of lunar exploration SMART-1 ⁶⁾ and the north-south control of geostationary satellites can be cited. It is conceivable that the Hall thruster will occupy a major position in future space promotion, and research and development is carried out in various research institute ⁷⁻⁹⁾ also in Japan. In addition, Japanese companies, IHI ¹⁰⁾ and Mitsubishi Electric ¹¹⁾ have been actively researching and developing and Hall thruster was adopted in Engineering Test Satellite No.9 as a main propulsion. ¹²⁾

As name showed, the shape and strength of the magnetic field have a great influence on its performance in Hall thruster. Moreover, the optimum magnetic field shape and optimum magnetic field strength not only depend on the individuality of the thruster and on the operating conditions such as discharge voltage and mass flow rate, but also depend on the integration operational time due to the change in acceleration channel shape. To make matters worse, the tendency of the changes depends on the individuality of the thruster and how to use the thruster. These characteristics makes predicted optimum control from the ground test beforehand or numerical simulation or theory to be difficult. In the current condition, operation of Hall thrusters is performed manually. In the near future, it can be expected that more Hall thrusters will be used, manual control procedures will be bankrupt due to the high cost in expense and human resources.

Therefore, intelligent power processing unit (PPU) is required for controlling Hall thruster automatically. As a RAIJIN project, we have been developing intelligent PPU that follows time-dependent optimum magnetic field configuration. As a first step, we adopt PID control ¹³⁾ and it showed successful results for controlling one condition. However, PID control cannot predict the optimal condition, another control procedures, neural network, should be introduced, since physics behind Hall thrusters is a nonlinear. Neural network is applied and utilized in a wide range of fields such as image recognition¹⁴⁾, stock price forecast and weather forecast ¹⁵⁾.

In this study, we have been developing thrust prediction code using neural network and verified this code compared with experimental results (training data). We also investigate the effect of initial value of weights and parameter updating methods on the accuracy of the prediction and calculation cost.

II. Neural Network

II-I. Neural network structure

In this study, we conducted supervised learning with 73 training data sets using back propagation ¹⁶⁾. As shown in Fig.1, we adjusted weights of the neural network in order to make thrust expressed by the network close to the target thrust when the operating conditions such as gas flow rate, discharge voltage and three coil currents which form the magnetic field are given.

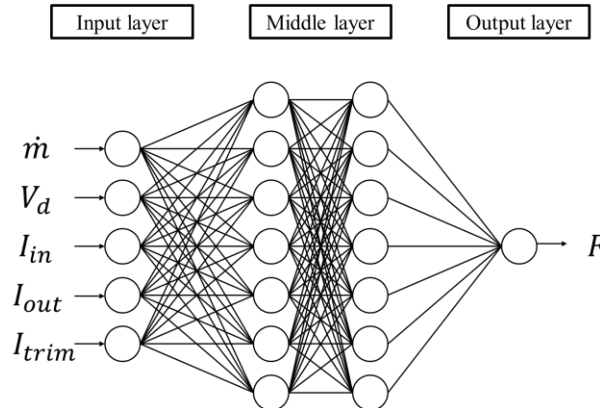


Figure.1 Neural network structure

As training data, we used 73 datasets of 5 kW anode layer type Hall thruster, RAIJIN94,¹⁷⁾ cooperatively developed by JAXA and 9 Universities. The experiment was conducted in a in an ISAS/JAXA ion engine endurance test vacuum chamber,⁴⁰⁾ 2 m diameter by 5 m length, evacuated by four cryogenic pumps (44,000 l/s for xenon), with the pressure kept below 6.6×10^{-3} Pa (for xenon) during thruster operation, with total mass flow rates of 13.5 mg/s(140 sccm) and a space science chamber in ISAS/JAXA, which was equipped with two cryopumps and one turbomolecular pump in a stainless-steel container of 2.5 m in a diameter and 5 m in length, the ultimate pressure was 2.7×10^{-5} Pa, and pressure was 5×10^{-3} Pa in operating condition (the thruster flow rate was 4.9 mg/s, cathode flow rate). The thrust is measured by a pendulum type thrust stand. The error of this thrust stand was 0.5 mN at 10 mN.

II-II. Update parameters

This time we used stochastic gradient descent (SGD) of back propagation (BP). SGD is optimization algorithm of machine learning. SGD can be written as equation (1).

$$\mathbf{W}^{(t+1)} = \mathbf{W}^{(t)} - \varepsilon \frac{\partial L}{\partial \mathbf{W}} \quad (1)$$

$\mathbf{W}$ is a parameter called the weight and this is updated by gradient of the loss function. L is loss function which indicates the error from the target value, and ε is learning rate. We used Rectified Linear Unit (ReLU) as the activation function of middle layers.

In neural network, the learning rate and weight initialization are important¹⁶⁾. If the learning rate is too small, it takes too much time for learning, and conversely, if it is too large it diverges and doesn't learn correctly. It is general recognition that success or failure in learning depends on what value is set as the initial value of weights. To improve learning accuracy and shorten learning period, we investigated the effect of initial value of weights and learning rate on them. In concrete terms, we used He initialization¹⁸⁾ as the initial value of middle layers and Ada Grad¹⁹⁾ as a method of learning rate decay.

He initialization is based on the 2015 paper by He et al., this is a method concerning an initial value of weight specialized for activation function of ReLU. In "He normalized initialization", weights are initialized with Gaussian distribution with standard deviation $\sqrt{2/n}$, where n is the number of previous nodes.

As an effective technique on the learning rate, there are some methods of gradually decaying the learning rate. Ada Grad is one of that methods and it creates custom-made values for each parameter. Weights are updated by following equation (3).

$$h^{(t+1)} = h^{(t)} + \frac{\partial L}{\partial \mathbf{W}} \odot \frac{\partial L}{\partial \mathbf{W}} \quad (2)$$

$$\mathbf{W}^{(t+1)} = \mathbf{W}^{(t)} - \varepsilon \frac{1}{\sqrt{h}} \frac{\partial L}{\partial \mathbf{W}} \quad (3)$$

$\mathbf{W}$ and L are as the same parameters as SGD. As shown in Equation 3, h holds the value of the gradient of the loss function as the sum of squares. At updating the parameter $\mathbf{W}$, the scale of learning is adjusted by multiplying by $1/\sqrt{h}$. Therefore, it enables to gradually decay learning rate for each element for each parameter.

III. Results

III-I. Initial value distribution effect

Figure 2 and 3 show the activation distribution (the output data distribution after assigned to ReLU) in each middle layers using Gaussian distribution with standard deviation of 0.01 and using He initialization (Gaussian distribution with standard deviation of $\sqrt{2/n}$, 0.166), respectively. We investigated the effect of initial value of weights on activations of middle layers. In these cases, we used the neural network with 5 layers and 100 nodes at each layer. Figure 2 shows activation in each layer was quite small value, even the first layer. Since quite small values flow on the neural network, the gradient of weights at back propagation becomes also small. Because of this, learning hardly proceeds. On the other hand, in case of He initialization, the spread of distribution in each layer is uniform as shown in Fig.3. The spread of data is kept uniform even if the number of layer increases, so it can be expected that appropriate values flow at back propagation. it means.

Figure 4(a) shows experimental thrust maps at propellant mass flow rate of 4.9 mg/s, the ratio of inner coil current to outer coil current of 1. Figure 4(b) and (c) show the reproduced thrust map deduced from the neural network whose initial weights distribution is Gaussian function with standard deviation of 0.01 and that He initialization, respectively. The reconstructed thrust by the neural network with Gaussian with standard deviation of 0.01 initialization is not in agreement with experimental thrust, on the other hand, that with He initialization is in good agreement with the

experimental one, considering the uncertainty within 5 %, as Figs 2 and 3 predicted. As a result, the neural network with 5 layers and 100 nodes with He initialization can reproduced target distribution of thrust sufficiently.

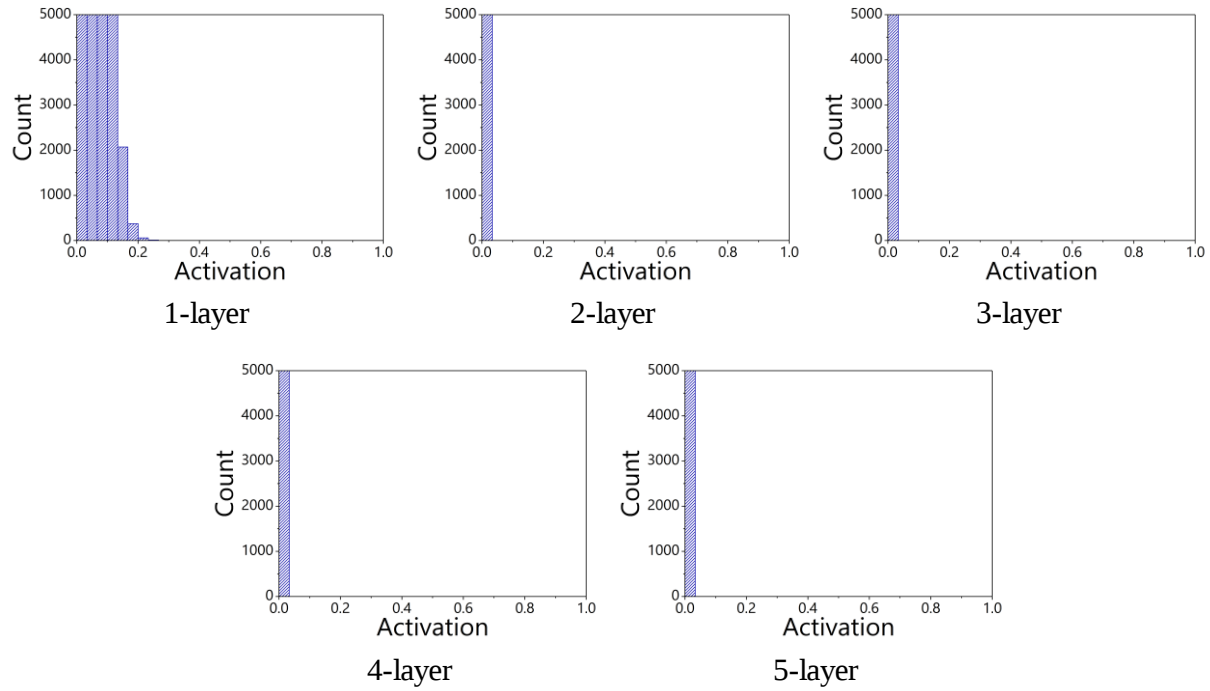


Figure. 2 Activation value of each layers using initial value with standard deviation of 0.01.

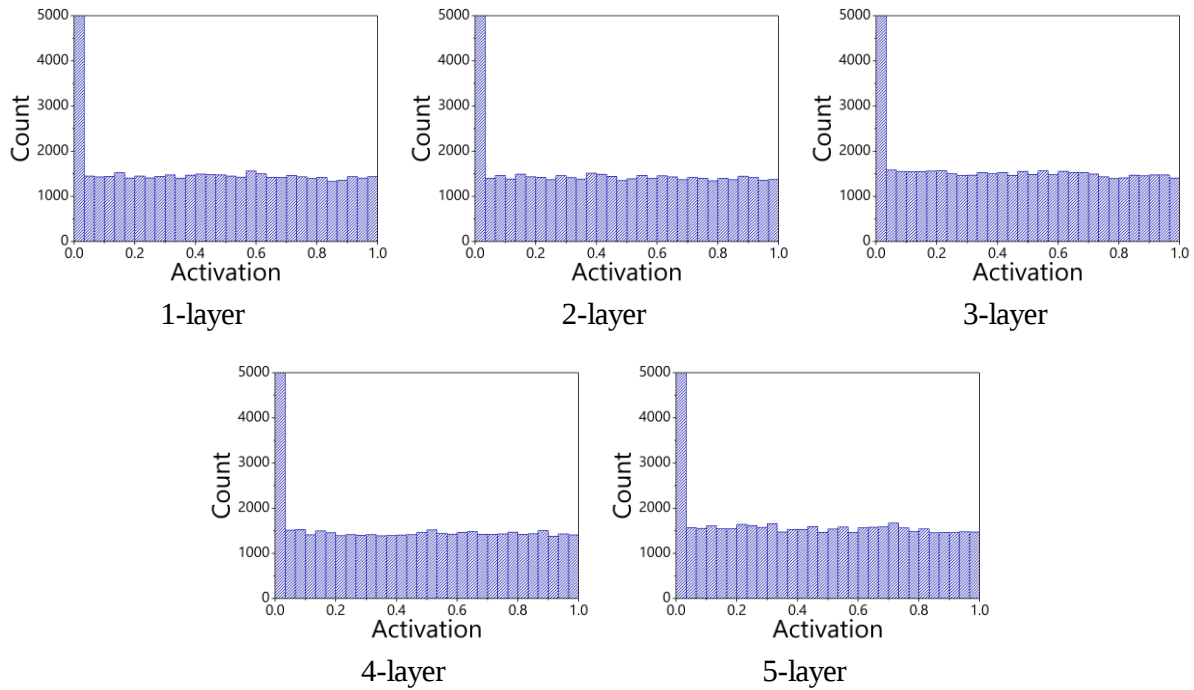


Figure. 3 Activation value of each layers using He initialization.

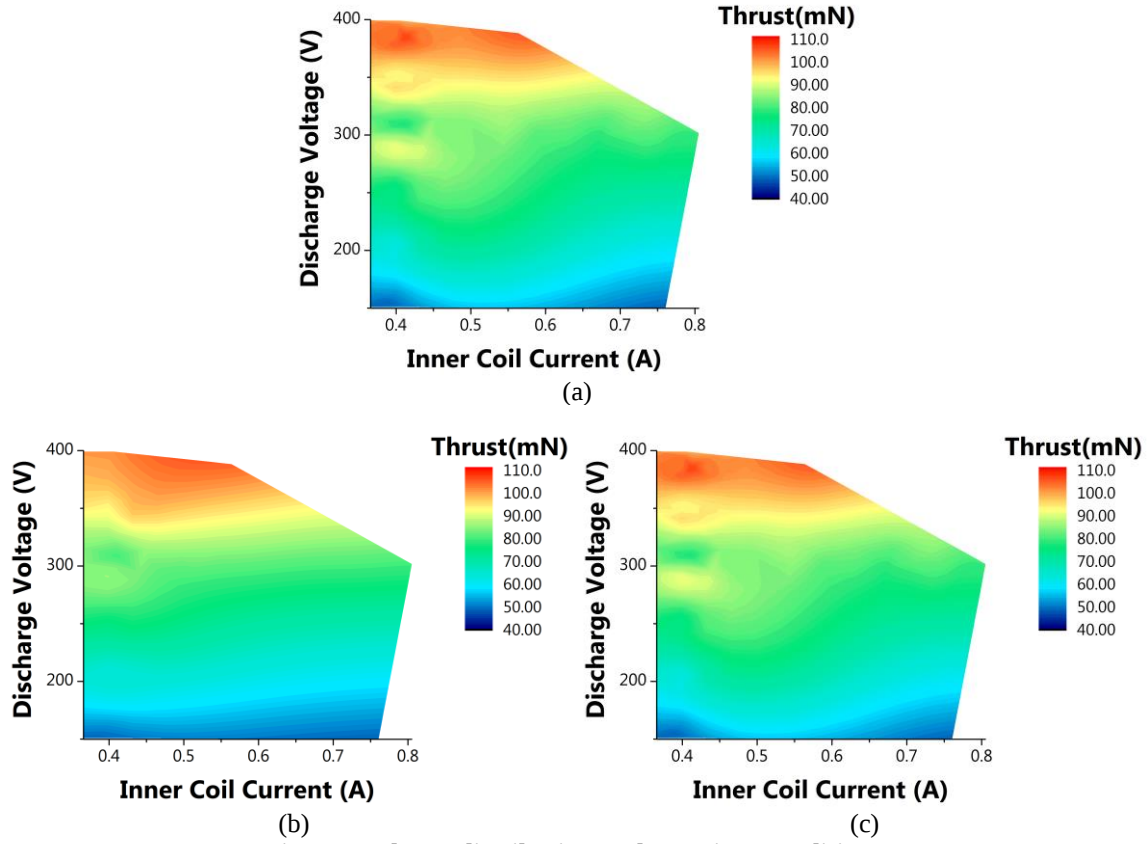


Figure. 4 Thrust distribution under various conditions,
(a) target thrust distribution, (b) reproduced thrust distribution with standard deviation of 0.01.
(c) reproduced thrust distribution with He initialization

III- II . Comparison of update methods

Figure 5 shows the comparison of learning speed. We examined that how the learning speed changes by introducing Ada Grad. The horizontal axis indicates number of iterations, and the vertical axis indicates the value of loss function. In this verification, we used the network with 5 layers and 100 nodes at each middle layer. We used He initialization as the initial value of weights in both SGD and Ada Grad. As shown in Fig.5, the adoption of Ada Grad improves the learning speed. We are now trying to the other updating parameters method, such as Momentum²⁰⁾, Adam²¹⁾, etc. Furthermore, since the result changes depending on the setting value of the hyper-parameter such as the learning rate, it is necessary to consider it.

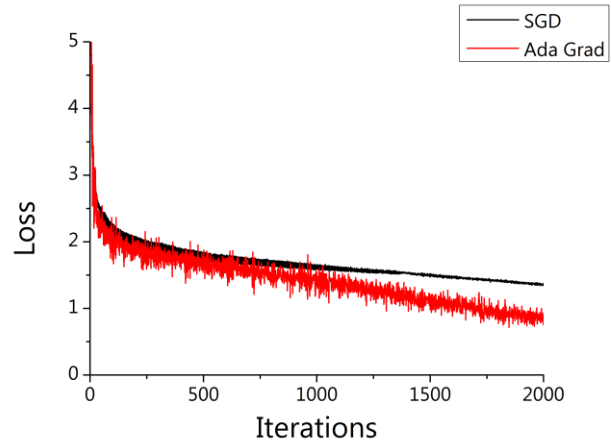


Figure 5. Comparison of update methods.

III-III. Prediction of thrust

Then we excluded 6 data with discharge voltage 150 V to 400 V from the training data, and made these data as unknown data. After learning with 67 data (73 – 6 data), we verified whether neural network could predict thrust of unknown data correctly. First, we used the network which has 5 layers with 100 nodes as middle layer. The result is shown in Fig. 6(a). The horizontal axis indicates discharge voltage, and the vertical axis indicates thrust. And the error bar indicates 5 % uncertainty from the thrust by experiment. The predicted thrust deviates from target thrust in range of discharge voltage more than 350 V. The following two things can be considered as this cause. First, since the degree of freedom that the network can express for the data is too large for this data set, this might result in overfitting. Second, the number of data in this region is small for the prediction of the thrust. So, we reduced the degree of freedom of network and confirmed whether overfitting occurred. That is, we changed the network structure to the network composed of 2 layers with 10 nodes. The result is shown in Fig. 6(b). Though the degree of freedom decreases, the prediction accuracy is improved compared with 5 layers and 100 nodes. It means the overfitting would be occurred.

The accuracy is still not enough, especially in range with discharge voltage is more than 350 V, this would be because the number of data is small or the degree of freedom is not enough. We have a plan to adopt a method such as dropout²²⁾ that suppresses overfitting even in the network with high expressivity as well as use more data sein order to improve the accuracy. Since it is not always possible to acquire many training data during actual operation, it is necessary to make network that can be predict the target thrust with few data sets.

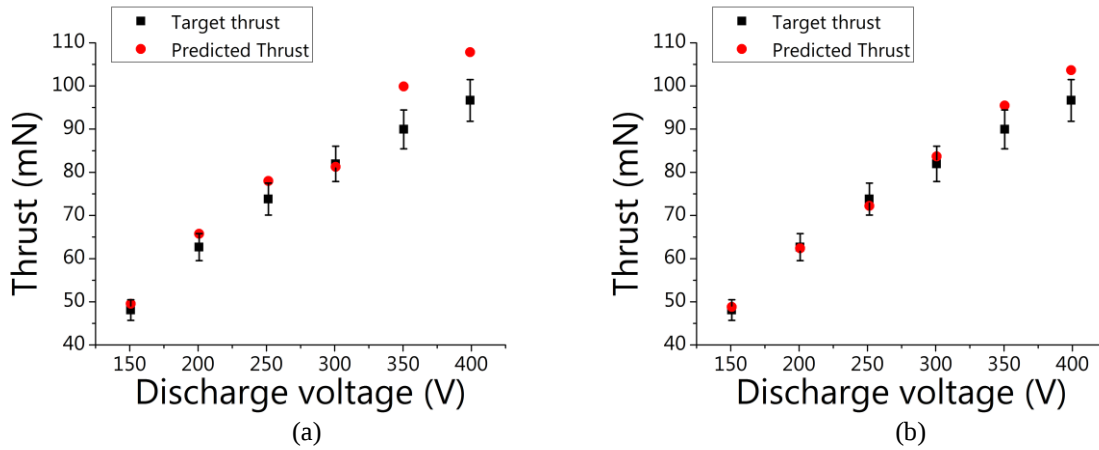


Figure 6. Prediction of unknown data by network (a) composed of 5 layers with 100 nodes, (b) composed of 2 layers with 10 nodes.

IV. Conclusion

In this study, we attempted to reconstruct and predict thrust of the Hall thruster using neural network. At that time, we investigated that how initial value of weights and parameter updating method affect learning. The results obtained are as follows.

- (1) By introducing He initialization, the activation of each layer became uniform.
- (2) Comparing result of SGD and Ada Grad, we could improve learning speed by introducing Ada Grad. But we should attempt other updating methods.
- (3) Prediction accuracy is poor in the network with high expressivity due to the small number of data.

In the future, we will employ the technique such as weight decay and dropout in order to proceed learning without overfitting even with small data. And we plan to construct the neural network that can automatically track parameters such as the magnetic field which must be applied to obtain target thrust.

Acknowledgments

The research results were obtained using Space Plasma Laboratory of ISAS, JAXA. This work was supported by JSPS KAKENHI Grant Number, JP 16K14506 and JP26289324, and Mitsubishi Electric Corporation and JAXA.

The contributions of Prof. Hitoshi Kuninaka, Prof. Kazutaka Nishiyama, Prof. Satoshi Hosoda, Prof. Ryudo Tsukizaki, Mr. Yushi Hamada and their staff are greatly appreciated.

References

- 1) <http://www.boeing.com/boeing/defense-space/space/bss/factsheets/702/702SP.page>
- 2) Yushi Hamada, Junhwi Bak, Rei Kawashima, Hiroyuki Koizumi, Kimiya Komuraaki, Naoji Yamamoto, Yusuke Egawa, Ikkoh Funaki, Shigeyasu Iihara, Shinatora Cho, Kenichi Kubota, Hiroki Watanabe, Kenji Fuchigami, Yosuke Tashiro, Yuya Takahata, Tetsuo Kakuma, Yusuke Furukubo, and Hirokazu Tahara, "Hall Thruster Development for Japanese Space Propulsion Programs", *Trans of JSASS*, Vol. **60**, (No. 5), 2017, pp. 320–326.
- 3) Kyoichi Kuriki, Yoshihiro Arakawa, "Introduction to Electric rocket "(in Japanese)The University of Tokyo press, 2003.
- 4) H. R. Kaufman, "Technology of Closed-Drift Thrusters," *AIAA journal*, Vol. **23**, No. 1, 1985, pp.78-86.
- 5) V. Kim, "Main physical feature and processes determining the performance of stationary plasma thrusters," *Journal of Propulsion and Power*, Vol. **14**, No. 5, 1998, pp. 736–743.
- 6) C. R. Koppel and D. Estublier, "The SMART-1 Electric Propulsion Subsystem," AIAA paper 2003-4545, 39th Joint Propulsion Conference, Huntsville, Alabama, July 20–23, 2003.
- 7) K. Komurasaki and Y. Arakawa, "Hall-Current Ion Thruster Performance," *Journal of Propulsion and Power*, Vol. 8, No. 6, 1992, pp. 1212–1216.
- 8) T. Miyasaka, Y. Shibata, K. Asatio and K. Segawa, "Investigation of Acceleration Channel Processes in Hall Thrusters by Particles Simulations," Thruster, Trans. of Japan Soc. for Aeronautical and Space Sci., Vol. 7, No. ists26 pp.Pb_83-Pb_88, 2009.
- 9) S. Yuge and H. Tahara, "Performance Enhancement of a 1-kW Anode-layer Hall Thruster," IEPC paper -2005-020, 2005.
- 10) Ikkoh Funaki,, Shigeyasu Iihara,Shinatora Cho, Kenichi Kubota , Watanabe Hiroki, Kenji Fuchigami, Yosuke Tashiro "Hall Thruster Development for All-electric Propulsion Satellite and Deep Space Explorers", AJCPP2016-047,2016.
- 11) T. Ozaki, Y. Inanaga, T. Nakagawa, Y. Kasai, K. Matsui. "Development status of high power xenon Hall thruster of MELCO," ISTSpaper 2006-b-34, 25th international symposium on space technology and science, 2006.
- 12) http://www.mext.go.jp/b_menu/shingi/gijyutu/gijyutu2/059/shiryo/_icsFiles/afieldfile/2017/05/22/1385842_3.pdf
- 13) Hiroyuki Osuga, Fjio Kurokawa, Naoji Yamamoto,Taichiro Tamida, A New Magnetic Flux Density Control Method to Improve Power Consumption of Hall Thruster, IEPC2011-182, 32nd International Electric Propulsion Conference, Wiesbaden , Germany, September, 2011.
- 14) Quoc Le, Marc'Aurelio Ranzato, Rajat Monga, Matthieu Devin, Kai Chen, Greg Corrado, Jeff Dean, Andrew Ng, " Building High-level Features Using Large Scale Unsupervised Learning", International Conference in Machine Learning (2012).
- 15) Veronica J. Berrocal • Adrian E. Raftery • Tilmann Gneiting • Richard C. Steed、 Probabilistic Weather Forecasting for Winter Road Maintenance、 American Statistical Association Journal of the American Statistical Association、 Vol. 105, No. 490 (2010)
- 16) Josh Patterson, Adam Gibson, "Deep Learning: A Practitioner's Approach," O'Reilly Media, 2017.
- 17) Yushi Hamada, Junhwi Bak, Rei Kawashima, Hiroyuki Koizumi, Kimiya Komurasaki, Naoji Yamamoto, Yusuke Egawa, Ikkoh Funaki, Shigeyasu Iihara, Shinatora Cho, Kenichi Kubota, Hiroki Watanabe, Kenji Fuchigami, Yosuke Tashiro, Yuya Takahata, Tetsuo Kakuma, Yusuke Furukubo, Hirokazu Tahara, "Hall Thruster Development for Japanese Space Propulsion Programs," *Trans of JSASS*, Vol. 60(2017) No. 5, p.320-326, <http://doi.org/10.2322/tjsass.60.320>
- 18) Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification, Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV) , Pages 1026-1034.
- 19) Duchi, John, Hazan, Elad, and Singer, Yoram. Adaptive subgradient methods for online learning and stochastic optimization, Journal of Machine Learning Research 12 (2011) 2121-2159.
- 20) Sutskever, Ilya, Martens, James, Dahl, George, and Hinton, Geoffrey. On the importance of initialization and momentum in deep learning. In Proceedings of the 30th International Conference on Machine Learning(ICML-

- 13), pp. 1139–1147, 2013.
- 21) Diederik P. Kingma, Jimmy Ba, Adam: A Method for Stochastic Optimization, conference paper at ICLR 2015.
 - 22) Nitish Srivastava , Georey Hinton , Alex Krizhevsky, Ilya Sutskever, Ruslan Salakhutdinov , Dropout: A Simple Way to Prevent Neural Networks from Overfitting Journal of Machine Learning Research 15 (2014) 1929-1958.