

Vision-based attitude estimation for spacecraft docking operation through deep learning algorithm

Thaweerath Phisannupawong[†], Patcharin Kamsing^{†*}, Peerapong Torteeka[‡], Soemsak Yooyen[†]

[†]Department of Aeronautical Engineering and Commercial Pilot, International Academy of Aviation Industry

King Mongkut's Institute of Technology Ladkrabang, Bangkok 10520, Thailand

[‡]National Astronomical Research Institute of Thailand, ChiangMai 50180, Thailand.

Email: thaweerath2009@gmail.com, insom.patcharin@gmail.com, peerapong@narit.or.th, soemsak.yo@kmitl.ac.th

Abstract—On-orbital services, especially in docking operation and other space object interaction. The missions need accurate, reliable, and robust detection to be an accurate in joining any interaction concerned. Two spacecrafts with an unknown mathematical model to predict the position and orientation, a computer vision-based attitude estimation system to detect the poses of spacecraft via camera is the key option of the mission.

In astronautics control, the position coordinates are normally represented as the cartesian coordinate system and used a quaternions coordinate system for orientation representation because quaternions can represent the orientation of spacecraft better than physical angle and can overcome the problem of singularity.

This paper aims to construct a model for both position and orientation estimation with public data. The input images are the dataset of Soyuz in the resolution of 1280x960, which is simulated by Unreal Engine 4. The implementation of this paper use GoogLeNet for a convolutional neural network model with the mathematical model of loss subject to direct regression.

The result shows that a position estimation is significantly accurate with having distance error smaller than 1 meter and tend to reduce when setting a proper scaling factor for loss function. The result demonstrates a high error for orientation estimation. However, the experiment expresses that both position and orientation estimation can be improved in case of selecting a suitable scaling factor of loss function.

Index Terms—Spacecraft Docking, On Orbit Services, Pose Estimation, Deep Learning, Convolutional Neural Network (CNN)

I. INTRODUCTION

Spacecraft position and attitude estimation are essential processes in a docking operation. In order to join two spacecrafts or other space interactions of multiple objects in space, the operations must have an accurate, reliable, and robust positioning and attitude determination system. For the unknown mathematical model of spacecraft equation of motion, the prediction of spacecraft attitude at any time cannot be determined by hand calculation or traditional programming. Thus, the deep learning algorithm can be used to solve the problem in this kind of situation. This means we only know the input and the output, which can match each other individually.

Spacecraft attitude estimation has been useful in several works for a cooperative system like the navigation or sensor on spacecraft [13], [14] and noncooperative in [7], [10]. In

this paper, the author mainly focuses on noncooperative detection on camera images processing through a deep learning algorithm.

The main idea of this paper is to construct a pre-trained model for the spacecraft positioning and attitude estimation system. By gathering the data as images of the spacecraft, Soyuz and relative positions and relative poses in quaternions of the spacecraft to the camera are separated in two sets for training and for model testing.

The primary condition required for deep learning is to have a massive amount of data provided in the training step. In outer space, the cost to conduct the data of images is too high to matched their position and attitude of the space object while the images are being taken. For the pre-train model implementation, it is an excellent choice to take an open data, which is already available. Those data will be used for the data analysis, through deep learning to construct the model, which can be used to train the additional data later.

The open-source data require to have a realistic appearance and be able to visualize the space appearance. The realistic background of the images and lighting conditions can be good samples for this pre-trained model development. The data generated using Unreal Engine 4 meets all the requirements of the high-quality data for the pre-trained model.

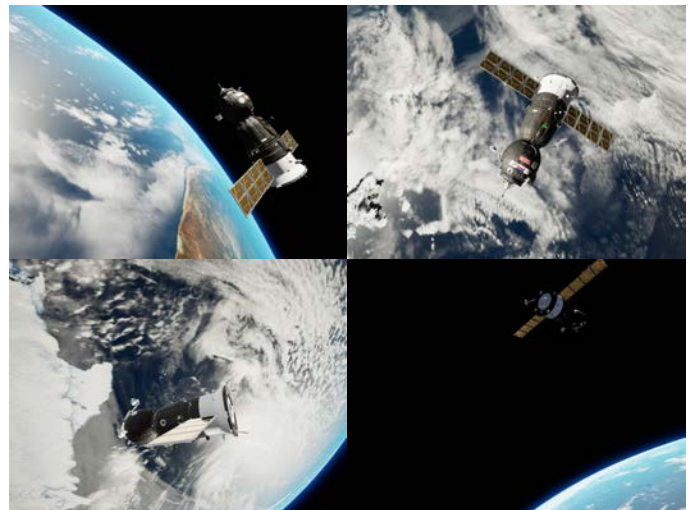


Fig.1 The images data generated from Unreal Engine 4

This work was supported by King Mongkut's Institute of Technology Ladkrabang, Bangkok 10520, Thailand.

The advantage of constructing the pre-trained model is to use the cheap data or open-source data for training the model in the said costly operation. The pre-trained model can reduce the cost of training with expensive data. In this case, they are real pictures captured from other spacecrafts to the targeted spacecraft, replaced by public data. The method of the pre-trained model constructs the pattern of neural network suitable for the operation. For this reason, it makes it flexible to the application in additional training with a few sets of images. Therefore, the cost is reduced.

The main concepts of the algorithm of pose estimation in many works are suitable examples in this operation. They demonstrate the method to construct a training model with many data taken. They presented a different method to construct the model for the neural network. For this paper, the public data generated from Unreal Engine 4 can replace the real data from the camera of spacecraft to the target object. Instead of relative poses of the camera in other application, this paper prefer to use the relative poses of the target with respect to the camera, to determine the pose of the target. However, in the term of images data and parameters of position and poses, they are the same in implementation.

This paper is demonstrated in the following section. Section II describes related works which is mainly the loss function in CNN. Section III is the implementation section beginning with dataset preparation, follow by the pose estimation algorithm construction, and the testing processes, respectively. Section IV is about the result and discussion. Conclusion is drawn from the final section of this paper.

II. RELATED WORKS

Regression in [3] derives the equation from the development of the training model. They demonstrate many methods for pose estimation for different methods and dimensions of rotation. In the case of a spacecraft as the rotating object, the implementation of spacecraft poses is beyond the determination via Euler angles. Thus, it requires quaternions to represent the angle of rotation of the object. Many provided public datasets have already provided containing the images those match the quaternions rotation values of those images individually.

According to the introduction, the dataset provided by [1] generated by URSO is the simulator built in Unreal Engine 4 to generate realistic images of the spacecraft surrounding by the environment of the outer space. The generated images are very realistic. They can visualize the condition of the outer space to the spacecraft in harsh lighting conditions and the realistic earth surface to be the background of the training and testing images. With provided positions and poses data. These data, which are similar to the real photo from the real space object, can be good examples for the model to be trained. They demonstrate the method of training with ResNet containing the pre-trained model as a CNN backbone. This method has a high accuracy and comes with a high complexity, which is hard and complicated to implement. In the same fashion, another work done on spacecraft is [9] the algorithm construct

through convolutional neural networks with the dataset of the PosE Estimation dataset(SPEED), which is the simulated data.

The approach in [2] takes a Convolutional Neural Network on the pose estimation of the camera relative to the scene to develop the algorithm that is practical in the application of pose estimation. They use the modified architecture of GoogLeNet convnet. They take the method of training, with interior data for less outdoor data to train the model. Several works through the Convolutional neural network, focus on the pose estimation of the object. They seem to be more applicable to this paper because the training of general object orientation is the same as the training of space objects. There are some works done on a single object in one image [4] and for multiple objects [5], [6], [8], [15]. In this paper, the author focuses on the pose estimation of a single object, which is a spacecraft. [11], [12] pose estimation frameworks are implemented on the Pascal dataset, which contain many kinds of objects for training in pose estimation field. Many methods demonstrated by many works can be applied to the training method of this paper. The implementation takes Unreal Engine 4 realistic images to do the pre-trained for the higher quality of data or real data of spacecraft or other space objects in the future.

From many previous works, they have shown the estimation processes suitable for their application. There are many methods of regression derived from many theories that can be implemented to the construction of deep neural networks. One of the primary considerations of this paper is the loss calculation to optimize the results and fine-tuning of scaling coefficients that can minimize the error in the results. The required method of loss calculation and scaling coefficients are different for each kind of considered object. The regression of spacecraft position and orientation will have its suitable method and scaling coefficient. Thus, the primary considerations are to select the most efficient CNN model, the mathematical model of calculation, and the scaling factor of calculation for the operation of the spacecraft.

Several works separate the terms of calculation in position loss and attitude loss. For position loss, most of the works implement Euclidean distance multiply with the scaling coefficient.

$$Loss_{position} = \beta_x \| x_i - x_{gt} \| \quad (1)$$

Where x_i is trial position vector from the layers of CNN and x_{gt} is the ground truth position vector which is informed in the dataset.

In general, Pose estimation works [1], [4], [5], [11] usually be employed by Geodesic distance with the equation.

$$Loss_{orientation} = \beta_q \arccos(|q_i^T q_{gt}|) \quad (2)$$

Where q_i is trial quaternions components from the layers of CNN and q_{gt} is the ground truth quaternions components which is informed in the dataset.

From equation (1) with (2), leads to the total loss equation

$$Loss_{total} = Loss_{position} + Loss_{orientation} \quad (3)$$

To minimize errors of estimation it need fine tuning of β_x and β_q that multiply separately in $Loss_{position}$ and $Loss_{orientation}$

III. IMPLEMENTATION

From the introduction, this paper aims to construct a pre-trained model for the position and attitude estimation of spacecraft. The implementation consist of three main sections, which are the dataset preparation, pose estimation algorithm construction, and testing, for those three implementations were trained on NVIDIA Geforce RTX 2080 on a laptop.

A. Dataset Preparation

In using the fundamental pre-trained model construction, [1] provided public data for practicing in the field of deep learning. This paper selects the dataset of Soyuz to construct the pre-trained model. The provided data, including 5000 images with position and quaternions, matched with the images individually of Soyuz separated into three sets which are train, test, and validate.

The appearance of images data including a spacecraft, the background of the earth's surface, black outer space appearance, and simulated directional light reflected on the spacecraft and the earth's surface. With the ability of Unreal Engine 4, it can simulate the realistic images and be able to display the relative position and angle between the object and its parent. Moreover, the angles can be converted to quaternions to overcome the problem of Euler matrix singularity.

Then, the prepared dataset for this paper condition is in the format of 1280x960 .png images separated to 4500 training images and 500 testing images with position and quaternions in the format of .csv

B. Pose Estimation Algorithm Construction

This paper employs a deep neural network for image classification, GoogLeNet, as the base architecture for the position and attitude regression. GoogLeNet contains modules called inception and 22 layers of convolutional neural network. In this paper, the preferred model is slightly different from ordinary GoogLeNet, switch to the modified GoogLeNet with 23 layers of the convolutional neural network as in [2]

The required input for this experiment is 224x224 images. Thus the preprocess of images before feeding to the model is the object to transform 1280x960 to smallest resolution as possible, which is 320x240. Then, it required to be cropped to 224x224. For this reason, the higher resolution than 320x240 will leads to insufficient image details, which were taken out via the cropping process.

Poses of spacecraft given in the public data include positions and poses of spacecraft. And they are in the format of seven column data set.

$$Poses = [X, Y, Z, q_1, q_2, q_3, q_4] \quad (4)$$

This seven dimension vector has to be sliced into.

$$Position = [X, Y, Z] \quad (5)$$

and

$$RotationAttitude = [q_1, q_2, q_3, q_4] \quad (6)$$

Where X, Y, Z is magnitude of distance in x,y and z axis respectively. For quaternions components q_1 represent the real part and q_2, q_3, q_4 is the component of vector part.

Another main consideration for algorithm development is loss function defining. The experiment apply equation (3) which combine loss from position and orientation. Equation (1) is the implementation for position loss function. The orientation loss is consider to be hard to get an accurate estimation. The experiment has been test by using equation(2) as orientation loss function. The result from equation (2) behave infinite at $|q_i^T q_{gt}| = 1$ to avoid that problem, the experiment employ the close form formulation of cosine of geodesic loss as equation (7)

$$Loss_{orientation} = \beta_q (1 - |q_i^T q_{gt}|) \quad (7)$$

But in the test of equation (7), iteration loss still result divert, Therefore, the experiment implements the natural exponential of geodesic loss in equation (8) as the loss function.

$$Loss_{orientation} = \beta_q \exp(1 - |q_i^T q_{gt}|) \quad (8)$$

Thus, $Loss_{total}$ in this experiment is defined as equation (9)

$$Loss_{total} = \beta_x \|x_i - x_{gt}\| + \beta_q \exp(1 - |q_i^T q_{gt}|) \quad (9)$$

Adam optimizer is used to minimize $Loss_{total}$ in each iteration during the development of the model. β_x and β_q can be tune by trial and error depending on dataset. Different datasets have different values of scaling coefficient.

C. Testing

Generally, the method of measuring the error, we mainly focus on the magnitude of the vector, or it can be focus on the size of the error. According to the equation (10),(11). The expression of loss which calculated by the median of magnitude error in each iteration of any dataset show as Fig(2). This expression loss is applied for the experiment.

$$Error_{position} = \|x_{predicted} - x_{actual}\| \quad (10)$$

and

$$Error_{rotation} = \|q_{predicted} - q_{actual}\| \quad (11)$$

And the errors in which show in this paper is the median of magnitude error in each test sample.

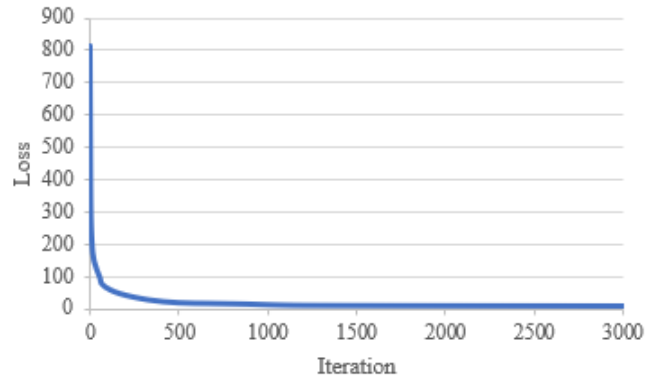


Fig.2 The relation of loss vs the number of iteration

In the experiment, at first, the iteration of fine-tuning done on 3000 repeated iteration of training. In different β_q and consider β_x is constant at 1. The reason for taking 3000 iteration value is 3000 iteration is the number of iteration, which the relation of loss versus iteration starts to become constant in its trend line. From Fig.2 as the typical deep learning at training procedure. At first time, the experiment ran the training at 10,000 iterations to see which value can be taken as a fine-tuning iteration number. Then, 3000 iteration is the number.

Then we start to test at the high value of β_q . To find the acceptable interval of β_q , we did the test on 3000 iterations at $\beta_q = [1,000],[100,000],[10,000,000]$. And pick some value between 100,000 and 10,000,000 and assume that this value can minimize the error in orientation. The result for rough tuning has shown in Fig.3 and can be seen that the point in which position and orientation are well-minimized lies in the interval of 10,000 to 100,000.

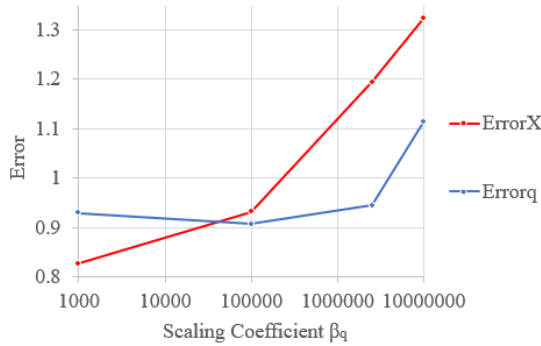


Fig.3 Approximated relation of Error versus scaling factor β_q while $\beta_x = 1$

TABLE I
APPROXIMATELY TUNING FOR β_q

β_x	β_q	Med.Position Error(m)	Med.Orientation Error
1	1000	0.82710419	0.92926225
1	100000	0.9316442	0.90732103
1	2547500	1.1937147	0.94482738
1	10000000	1.3219991	1.11328518

After this procedure, The result of β_q at 100,000 is already acceptable. Then we convert the error that obtained in quaternions system to be seen clearly in more physical terms of Euler Angle [16], [17].

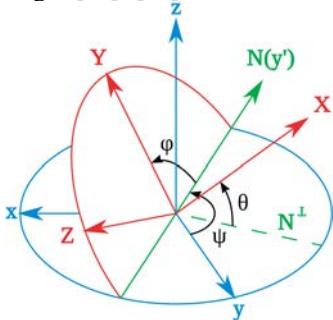


Fig.4 Euler Angles Representation

$$EulerAngles = [\phi, \theta, \psi] \quad (12)$$

$$Angle_x = \phi = \arctan \left[\frac{2(q_1q_2 + q_3q_4)}{(1 - 2(q_2^2 + q_3^2))} \right] \quad (13)$$

$$Angle_y = \theta = \arcsin [2(q_1q_3 + q_4q_2)] \quad (14)$$

$$Angle_z = \psi = \arctan \left[\frac{2(q_1q_4 + q_2q_3)}{(1 - 2(q_3^2 + q_4^2))} \right] \quad (15)$$

Where ϕ, θ, ψ is the angle of rotating frame around x,y, and z axis respectively. The equations, (13),(14),(15) apply the quaternions component as the representation in equation (6).

For the implementation in spacecraft dynamics control, quaternions representation is better because it does not have any problems with rotation sequences [18], [19] as the Euler angles. However, the Euler angles which convert from quaternions system to Euler angles, will lie on the same sequence that can be compared in the angular distance to visualize more physically.

IV. RESULT AND DISCUSSION

From the section III, the final training model given by $\beta_q=100,000$ is acceptable for approximate tuning. The experiment set loss function as equation(9), $\beta_x = 1, \beta_q = 100,000$, learning rate at 0.0001 and trained on 3,000 iteration via GoogLeNet and Adam optimizer.

Tbale II shows the results in the form of median orientation error at $\beta_q = 100,000$

TABLE II
RESULT OF POSITION AND OREINTATION IN QUARTEINION

β_x	β_q	Med.Position Error(m)	Med.Orientation Error
1	100000	0.9316442	0.90732103

After the manipulation in equation (13),(14),(15) to transform to the angles around three axis of Euler Angles. It demonstrate the output as Table III.

TABLE III
RESULT OF ORIENTATION IN EULER ANGLES REPRESENTATION

β_q	Med.Error ϕ (deg)	Med.Error θ (deg)	Med.Error ψ (deg)
100000	91.74883037	30.10585343	84.42547416

The overall results show that the orientation error is high even though in quaternions or Euler angles system and should continue to conduct further study to reduce such an error to achieve the maximum accuracy. However, the experiment can obtain a reasonable result of position error.

V. CONCLUSION

The paper aims to construct the pre-trained model for spacecraft position and attitude determination from 2D images, and from available data simulated based on the real outer space appearance — algorithm construction on GoogLeNet with Adam optimizer. The result shown above is not accurate

enough for orientation but it work well with position estimation.

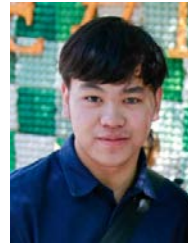
However, the method demonstrates in this paper can be reproduced. It needs more fine-tuning of the scaling coefficient and optimized other parameters concerned with reducing error such as the learning rate, the number of iteration, or implementing other optimization algorithm to work together to improve overall performance to be more precise and accurate making it practical to the real docking operation and other space interaction missions in the future.

VI. ACKNOWLEDGMENT

The author would like to thanks all staff and teachers from the International Academy of Aviation Industry, King Mongkut's Institute of Technology Ladkrabang and the National Astronomical Research Institute of Thailand. Moreover, the author is grateful for the advisers of this project for giving a lot of useful advice and solutions to the problems to complete this project.

REFERENCES

- [1] Pedro F. Proença, and Yang Gao, "Deep Learning for Spacecraft Pose Estimation from Photorealistic Rendering", ArXiv, arXiv:1907.04298, 2019
- [2] Alex Kendall, Matthew Grimes, and Roberto Cipolla, "PoseNet: A Convolutional Network for Real-Time 6-DOF Camera Relocalization", in IEEE International Conference on Computer Vision (ICCV), pp2938-2946, 2015
- [3] Shuai Liao, Efstratios Gavves, and Cees G. M. Snoek, "Spherical Regression: Learning Viewpoints, Surface Normals and 3D Rotations on n-Spheres", ArXiv, arXiv:1904.05404, 2019
- [4] Seyed Sadegh Mohseni Salehi, Shadab Khan, Deniz Erdogmus, and Ali Gholipour, "Real-time Deep Pose Estimation with Geodesic Loss for Image-to-Template Rigid Registration", IEEE Transactions on Medical Imaging PP(99):1-1 DOI: 10.1109/TMI.2018.2866442
- [5] Jimmy Wu, Bolei Zhou, Rebecca Russell, Vincent Kee, Syler Wagner, Mitchell Hebert, Antonio Torralba, and David M.S. Johnson, "Real-Time Object Pose Estimation with Pose Interpreter Networks", 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2018)
- [6] Thanh-Toan Do, Ming Cai, Trung Pham, Ian Reid, "Deep-6DPose: Recovering 6D Object Pose from a Single RGB Image", ArXiv, arXiv:1802.10367, 2018
- [7] Renato Volpe, Marco Sabatini, Giovanni and, B. Palmerini, "Pose and Shape Reconstruction of a Noncooperative Spacecraft Using Camera and Range Measurements", International Journal of Aerospace Engineering 2017:1-13
- [8] Yu Xiang, Tanner Schmidt, Venkatraman Narayanan, and Dieter Fox, "PoseCNN: A Convolutional Neural Network for 6D Object Pose Estimation in Cluttered Scenes", Robotics: Science and Systems 2017 DOI:10.15607/rss.2018.xiv.019
- [9] Sumant Sharma, Connor Beierle, and Simone D'Amico, "Pose estimation for non-cooperative spacecraft rendezvous using convolutional neural networks" IEEE, AIAA/AAS Space Flight Mechanics Meeting January 2019
- [10] Sumant Sharma, Jacopo Ventura, Simone, and D'Amico, "Robust Model-Based Monocular Pose Initialization for Noncooperative Spacecraft Rendezvous" Journal of Spacecraft and Rockets 55(1):1-16 DOI: 10.2514/1.A.34124, 2018
- [11] Siddharth Mahendran, Haider Ali, and René Vidal, "3D Pose Regression Using Convolutional Neural Networks", 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)
- [12] Siddharth Mahendran, Haider Ali, and René Vidal, "A Mixed Classification-Regression Framework for 3D Pose Estimation from 2D Images", BMVC 2018
- [13] J.M. Kelsey, Jeffrey Byrne, M. Cosgrove, Show, and Raman K. Mehra, "Vision-based relative pose estimation for autonomous rendezvous and docking", IEEE Aerospace Conference Proceedings 2006:20 pp. DOI: 10.1109/AERO.2006.1655916
- [14] Brent Edward Tweddle, "Computer vision based navigation for spacecraft proximity operations", Massachusetts Institute of Technology, 2010
- [15] Bugra Tekin, Sudipta Sinha, and Pascal Fua, "Real-Time Seamless Single Shot 6D Object Pose Prediction", 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)
- [16] Wikipedia - "Euler angles", https://en.wikipedia.org/wiki/Euler_angles
- [17] Jack B. Kuipers, "Quaternions And Rotation Sequence", Princeton University Press, Princeton, New Jersey, 1998
- [18] Anton H.J. de Ruiter, Christopher J. Damaren, and James R. Forbes, "Spacecraft Dynamics and Control an Introduction", Wiley, 2013
- [19] Dipl.-Ing. Karsten Groÿekathöfer, and Dr. -Ing. Zizung Yoon, "Introduction into quaternions for spacecraft attitude representation", Technical University of Berlin, 2012



Thaweerath Phisannupawong is currently a bachelor degree student, major in Aeronautical Engineering and Commercial Pilot (International Program), from the Department of Aeronautical Engineering, International Academy of Aviation Industry. His research interests are included the image processing and application of deep learning on aeronautical engineering.



Patcharin Kamsing received doctoral degree from the University of Chinese Academy of Sciences, Beijing, China sponsor by CAS-TWAS president fellowship. She currently is lecturer and assistant dean of International Academy of Aviation Industry, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand. Her research is image processing and remote sensing applications.



Peerapong Torteeka received doctoral degree in Celestial Mechanics and Applied Astrometry Engineering at National Astronomical Observatories of Chinese Academy of Sciences (NAOC), University of Chinese Academy of Sciences, Beijing, China. He currently works at National Astronomical Research Institute of Thailand, research interests include passive optical-based small and dim space debris recognition and tracking system, moving object extraction, autonomous detection, robotics telescope for the space observation.



Soemsak Yooyen currently is Dean of International Academy of Aviation Industry, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand. He received Doctor of Philosophy from Mie University, Japan. His research is about materials science.