# Tennis Player Training Support System based on Sport Vision

Kohei Arai[1], Toshiki Nishimura[2], Hiroshi Okumura[3]
Dept. of Information Science
Saga University
Saga City, Japan

***Abstract*—Sports vision based tennis player training support system is proposed. In sports, gaze, dynamic visual acuity, eye movement and viewing place are important. In sports vision, Static eyesight, Dynamic visual acuity, Contrast sensitivity, Eye movement, Deep vision, Instant vision, Cooperative action of eye, hand and foot, and Peripheral field are have to be treated. In particular for the tennis, all of the items are very important. Furthermore, trajectory of gaze location and tennis racket stroke gives some instructions for skill-up of tennis play. Therefore, sports vision based tennis player training system is proposed. Through experiment, it is found that the proposed system does work well for improvement of tennis players' skills.**

***Keywords*—Sport vision; static eyesight; dynamic visual acuity; contrast sensitivity; eye movement; deep vision; instant vision; cooperative action of eye; hand and foot; peripheral field***

## I. Introduction

Sports vision based tennis player training system is proposed. There is Sport Vision Association in Japan[1]. In sports, gaze, dynamic visual acuity, eye movement and viewing place are important. In sports vision, Static eyesight, Dynamic visual acuity, Contrast sensitivity, Eye movement, Deep vision, Instant vision, Cooperative action of eye, hand and foot, and Peripheral field are have to be treated. In particular for the tennis, all of the items are very important.

Face detection from the acquired imagery data is common and popular technology. Computer input system based on viewing vector estimation with iris center detection from face image acquired with web camera allowing users' movement is proposed [1]. Also, method for face identification with Facial Action Coding System: FACS based on eigen value decomposition is proposed [2]. The first book of its kind devoted to the emerging field of computer vision in sports is published [3]. Quite recently, the 5th International Workshop on Computer Vision in Sports (CVsports) at CVPR 2019 is held in Long Beach California, U.S.A [4]

A new Olympic vault dataset is proposed [5] and present three frameworks for action quality assessment which improve upon published results: C3D-SVR, C3D-LSTM and C3D-LSTM-SVR.The frameworks mainly differ in the way they aggregate clip-level C3D features to get a video-level description. This video-level description is expressive about the quality of the action. The task of detecting swimming strokes in the wild is demonstrated [6]. However, without modifying the model architecture or training method, the process is also shown to work equally well on detecting tennis strokes, implying that this is a general process. The outputs of the system are surprisingly smooth signals that predict an arbitrary event at least as accurately as humans (manually evaluated from a sample of negative results).

A convolutional neural network (CNN) has been designed to interpret player actions in ice hockey video [7]. The hourglass network is employed as the base to generate player pose estimation and layers are added to this network to produce action recognition. As such, the unified architecture is referred to as action recognition hourglass network, or ARHN. ARHN has three components. Group activity recognition in sports is often challenging due to the complex dynamics and interaction among the players. In this paper, we propose a recurrent neural network to classify puck possession events in ice hockey. Our method extracts features from the whole frame and appearances of the players using a pre-trained convolutional neural network. In this way, our model captures the context information, individual attributes and interaction among the players [8].

In order for action recognition to be useful in sports analytics a finer-grained action classification is needed. For this reason we focus on the fine-grained action recognition in tennis and explore the capabilities of deep neural networks for this task. In our model, videos are represented as sequences of features, extracted using the well-known Inception neural network, trained on an independent dataset. Then a 3-layered LSTM network is trained for the classification [9]. A learning-based framework that takes steps towards assessing how well people perform actions in videos is proposed [10]. The approach works by training a regression model from spatiotemporal pose features to scores obtained from expert judges. Moreover, the approach can provide interpretable feedback on how people can improve their action.

Using the line-of-sight camera, acquire eye-gaze video and analyze the difference between the beginner and the experienced person or the gaze at good and bad times. Also, I would like to support technical improvement by proposed system. Using head set of work camera, players' gaze is estimated together with ball trajectory for prediction. Then, some instructions can be provided to the player in concern. This is what I intend to do. Moreover, it is possible to make a simulated experience by Virtual Reality: VR from the line-of-sight image.

[1] http://www.sports-vision.jp/deta.htm

In the next section, the proposed system for improving tennis players' skills is described. Then, preliminary experiment is followed. After that, conclusion is described together with some discussions. Finally, future works are followed.

## II. Proposed System

### A. Design Concept

The proposed system provides expected position of tennis ball immediately after the ball is hit by the offensive player, as well as instructions and week points to tennis players. Fig. 1 shows design concept of the proposed system. There are two cameras, one is observe the tennis court and the other one is attached to the tennis player's head in the system.

From the video data acquired #1 camera, tennis ball is looked at and always calculate the trajectory and predict the expected tennis ball position touch down on the court. Therefore, instruction can be provided to the tennis player. Thus, the tennis player can take a next action so quickly. On the other hand, #2 camera is always looking forward forehead direction, head pose direction. Therefore, it is understand which direction is tennis player is looking at. Thus, week points can be provided when the tennis player is not looking at the appropriate direction after the play. That will help to improve their skills.

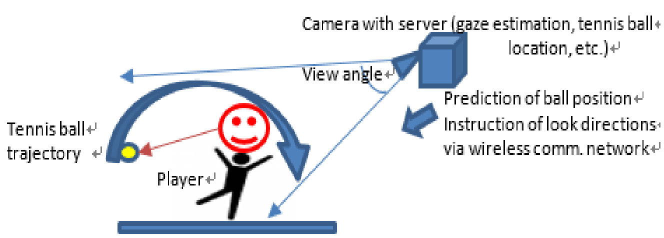



Fig 1. Design Concept

Tennis player wears GoPro of camera and play. Tennis ball can be extracted from the acquired moving picture using OpenCV. The trajectory of the extracted tennis ball and gaze locations are plotted and displayed onto computer screen. Then, valuable instructions are given to the players.

### B. Software and Hardware

Tennis ball can be extracted from the video and analyze where the players are looking together with gaze direction. There are some actions in tennis plays, Serve, Stroke, Volley, Smash, Match (singles). These actions are to be identified from the acquired videos. The identified action types can be provided to the tennis player through wireless communication network.

Major hardware and software used are as follows:

- GoPro HERO6 (Camera)
- OpenCV 3.1.0 for image analysis
- Python 2.7 or Visual Studio2017 C++
- Major specification of GoPro HERO6 is as follows,
- Weight: 117 g

Action camera

Image stabilization

Hi-Vision

Waterproof function

Field of view: "Wide angle", "fisheye"

Frame rate 4K / 60fps, 1080p / 240fps

Fig. 2 shows outlook of the GoPro HERO6.

(a)Back view (b)Front view

Fig 2. Outlook of the GoPro HERO6.

### C. Major Characteristics of Tennis Player

From data by SONY smart tennis sensor[2] (SONYSmartTennisSensor), major characteristics of tennis player are as follows:

Male average speed in 20's

Forehand stroke: 90 km / h (25 m / s)

Serve: 115 km / h (31 m / s)

The vertical length of the tennis court (the distance between the players): 23.77 m (about 25 m)

Until the hit ball reaches the opponent

Forehand stroke: about 1.0 second

Serve: about 0.8 seconds

### D. OpenCV

In order to analyze the acquired imagery data, OpenCV is used. The main functions that can be done are as follows:

- Filtering
- Matrix operation
- Object tracking
- Area segmentation
- Camera calibration
- Feature point extraction
- Object Recognition
- GUI
- Machine learning
- Panorama synthesis

[2] http://tennisblog.smartsports.sony.net/entry/worlddata-japan-detail-stroke

- Computational photography

Color based method is used for tennis ball detection. Detection of tennis ball is performed not only by using circle detection but also by color extraction as follows:

*1)* Hue, Saturation, Value: HSV conversion of the original image

*2)* Extract only yellow color

*3)* Extraction of contour (extraction of maximum contour only)

*4)* Draw the minimum circumscribed circle

Fig. 3 shows examples of detected circles using color information based method. The proposed tennis ball detecting method is based on the combination between shape based and color based methods. Therefore, tennis ball can be detected even if the tennis ball is not circle shape and even if the tennis ball is not yellow color.


Fig 3. Example of yellow colored object detection.

### E. *Tennis Ball and Gaze Location Trajectories as well as the Acquired Image of the One Forehand Stroke*

Tennis ball and gaze location trajectories can be analyzed in the 1080 by 1920 pixels of computer screen from the acquired moving pictures. One of the examples is shown in Fig. 4. Also, tennis ball location trajectory is analyzed and displayed onto computer screen as shown in Fig. 5. Furthermore, Fig. 6(a) and (b) show the acquired image of the one forehand stoke image of an expert and that of a beginner.

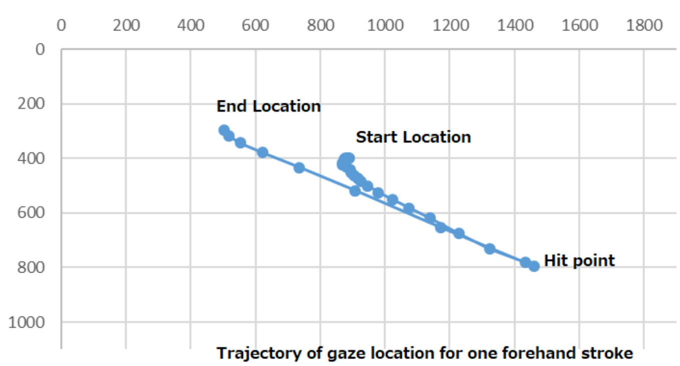



Fig 4. Example of gaze location trajectory for one forehand stroke.

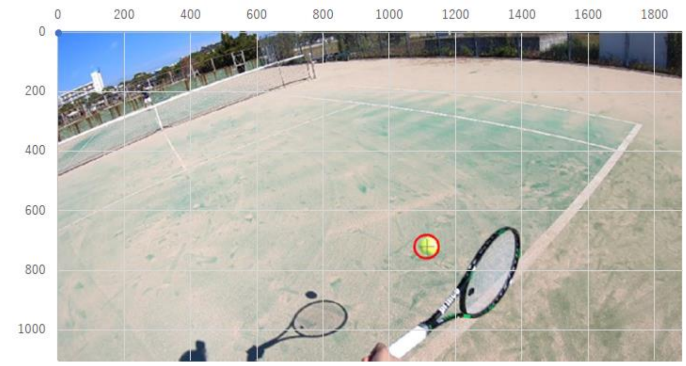



Fig 5. Example of tennis ball location trajectory.

(a)Expert

(b)Beginner

Fig 6. Acquired images with GoPro camera attached to an expert and a beginner of tennis players.

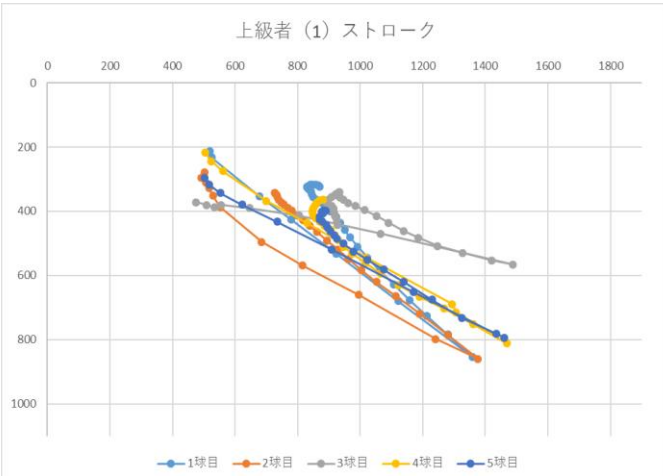
Thus, valuable instructions can be made available for tennis players based on the trajectories of ball and gaze locations.

## III. Experiment

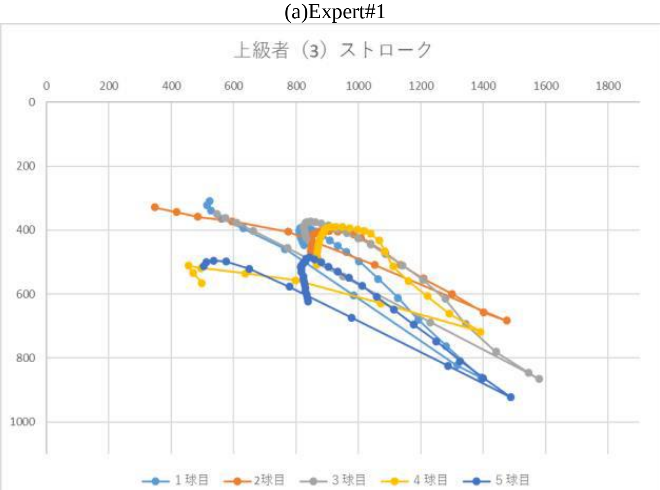
Six experts and six beginners of tennis players are participated to the experiment. Trajectories of the gaze location are analyzed for five forehand strokes. Fig. 7(a) to (f) shows the trajectories of experts while Fig. 7(g) to (l) shows those of beginners.

Where frame rate is 59 f/s which corresponds to about 0.0179 sec. per frame while time duration between two dots in the gaze trajectory figure is 0.0179 sec. From the trajectories, hitting point and stroke position and its blurring are recognized. It is quite obvious that head pose is stable, hitting point is also stable, pulling jaw, tennis ball is caught with eyes for experts while opposite for beginners. In other word from the sport vision of point of view, experts are good at the followings:

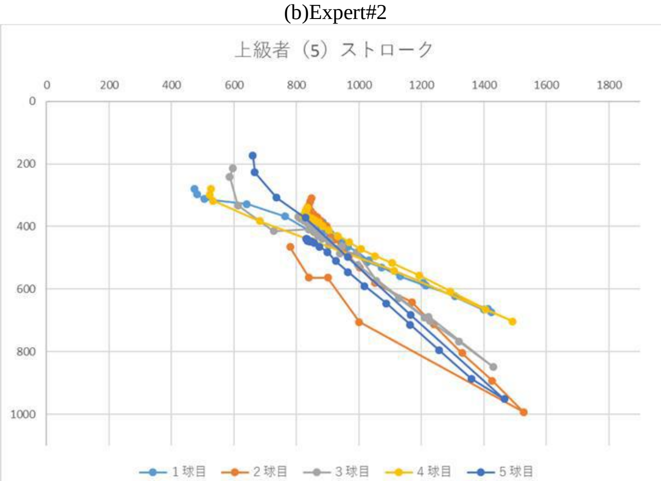
- Peripheral vision
- kva motion vision
- Deep vision
- Fixed gaze



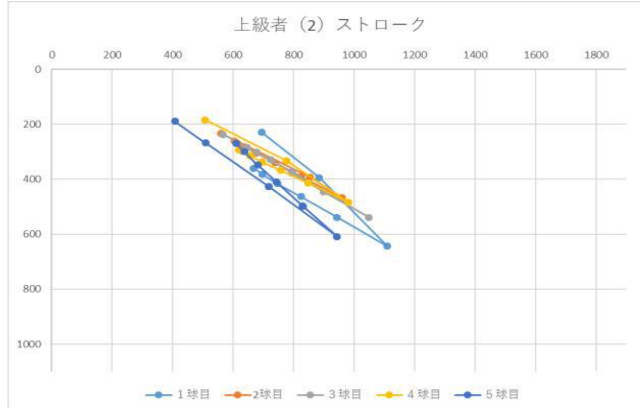

(a)Expert#1



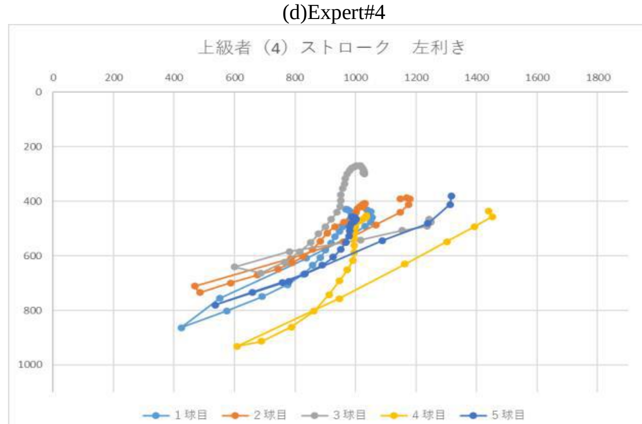

(b)Expert#2



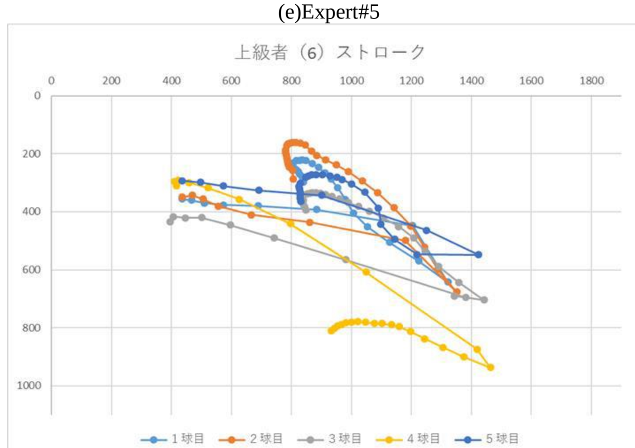

(c)Expert#3



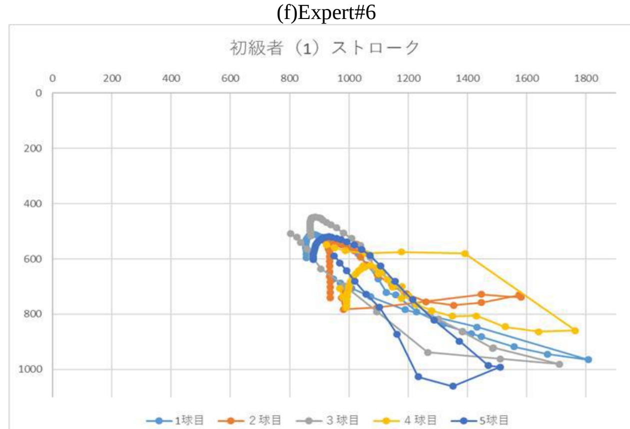

(d)Expert#4



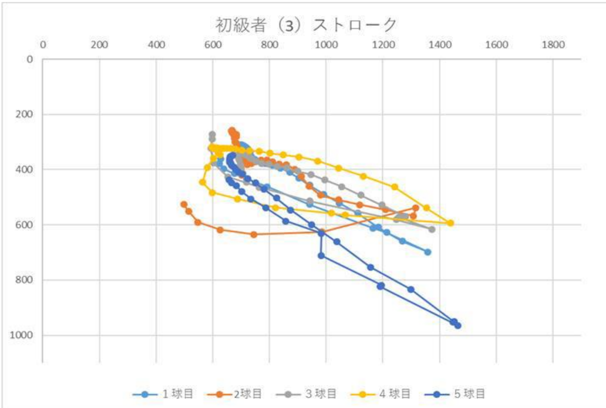

(e)Expert#5




(f)Expert#6

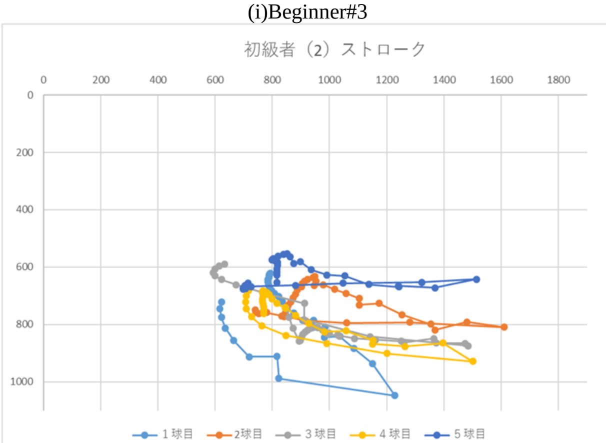



(g)Beginnert#1

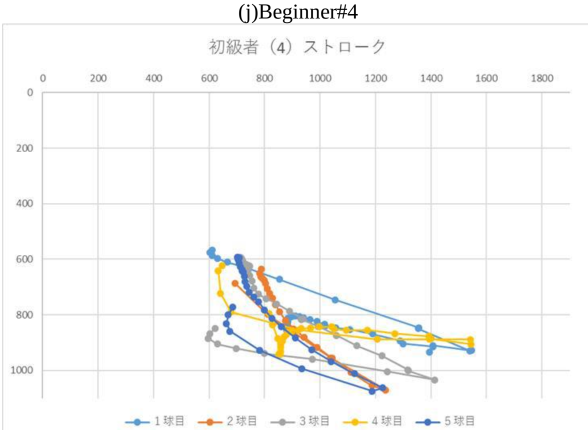



(h)Beginner#2

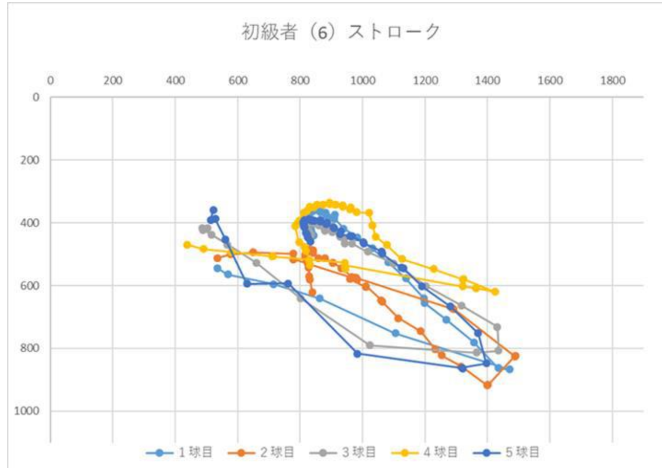

(i)Beginner#3




(j)Beginner#4




(k)Beginner#5




(l)Beginner#6

Fig 7. Gaze trajectories for 5 forehand strokes of experts and beginners.

Also, expert views opponent, catches the entire tennis court, catches the tennis ball properly as shown in Fig. 8(a) while beginner cannot view opponent, does not catch tennis ball properly, does not catch tennis court as shown in Fig. 8(b). This is the results from the acquired image of experts and beginners at the hitting point.

(a)Expert

(b)Beginner

Fig 8. Acquired image of experts and beginners at the hitting point.

## IV. Conclusion

Sports vision based tennis player training support system is proposed. In sports, gaze, dynamic visual acuity, eye movement and viewing place are important. In sports vision, Static eyesight, Dynamic visual acuity, Contrast sensitivity, Eye movement, Deep vision, Instant vision, Cooperative action of eye, hand and foot, and Peripheral field are have to be treated. In particular for the tennis, all of the items are very important. Furthermore, trajectory of gaze location and tennis racket stroke gives some instructions for skill-up of tennis play. Therefore, sports vision based tennis player training system is proposed. Through experiment, it is found that the proposed system does work well for improvement of tennis players' skills.

It is found that expert views opponent, catches the entire tennis court, catches the tennis ball properly while beginner cannot view opponent, does not catch tennis ball properly, does not catch tennis court

For future research works, real time instruction has to make available for tennis players.

*1)* Points looking at for expertized tennis player. This fact will be clarified.

*2)* Then, action classification from the acquired moving pictures will be conducted. If it is possible to identified just after the initial action, then next action can be recognized. After that, some preparatory action can be done with the recognized actions.

## Acknowledgment

The author would like to thank Prof. Dr. Osamu Fukuda in the Information Science Department of Saga University for his valuable comments and suggestions.

## References

[1] Kohei Arai and Hiromi Uwataki, Computer input system based on viewing vector estimation with iris center detection from face image acquired with web camera allowing users' movement, Electronics and Communication in Japan, 92, 5, 31-40, John Wiley and Sons Inc.,2009.

[2] Kohei Arai, Method for face identification with Facial Action Coding System: FACS based on eigen value decomposition, International Journal of Advanced Research in Artificial Intelligence, 1, 9, 34-38, 2012.

[3] https://www.springer.com/us/book/9783319093956

[4] http://www.vap.aau.dk/cvsports/

[5] https://arxiv.org/pdf/1611.05125.pdf

[6] http://homepage.cs.latrobe.edu.au/zhe/files/StrokeDetection.pdf

[7] http://openaccess.thecvf.com/content_cvpr_2017_workshops/w2/papers/Fani_Hockey_Action_Recognition_CVPR_2017_paper.pdf

[8] http://openaccess.thecvf.com/content_cvpr_2017_workshops/w2/papers/Tora_Classification_of_Puck_CVPR_2017_paper.pdf

[9] http://www.doc.ic.ac.uk/~wjk/publications/vinyes-knottenbelt-cvsports-2017.pdf

[10] https://www.csee.umbc.edu/~hpirsiav/quality.html

## Authors Profile

**Kohei Arai,** He received BS, MS and PhD degrees in 1972, 1974 and 1982, respectively. He was with The Institute for Industrial Science and Technology of the University of Tokyo from April 1974 to December 1978 and also was with National Space Development Agency of Japan from January, 1979 to March, 1990. During from 1985 to 1987, he was with Canada Centre for Remote Sensing as a Post Doctoral Fellow of National Science and Engineering Research Council of Canada. He moved to Saga University as a Professor in Department of Information Science on April 1990. He was a councilor for the Aeronautics and Space related to the Technology Committee of the Ministry of Science and Technology during from 1998 to 2000. He was a councilor of Saga University for 2002 and 2003. He also was an executive councilor for the Remote Sensing Society of Japan for 2003 to 2005. He is an Adjunct Professor of University of Arizona, USA since 1998. He also is Vice Chairman of the Commission-A of ICSU/COSPAR since 2008. He received Science and Engineering Award of the year 2014 from the minister of the ministry of Science Education of Japan and also received the Bset Paper Award of the year 2012 of IJACSA from Science and Information Organization: SAI. In 2016, he also received Vikram Sarabhai Medal of ICSU/COSPAR and also received 20 awards. He wrote 34 books and published 520 journal papers. He is Editor-in-Chief of International Journal of Advanced Computer Science and Applications as well as International Journal of Intelligent Systsems and Applications. http://teagis.ip.is.saga-u.ac.jp/