Digital Life
The Conference of Digital Life vol.2

Proceedings Presentation

A-1. Discussion on Detection of Dangerous Objects on the Road Surface Leading to Motorcycle Accidents Using a 360-degree Camera
Haruka Inoue and Yuma Nakasuji (Osaka University of Economics)

A-2. Fundamental Research on Class Judgment of Preschooler in Early Childhood Education
Tatsuma Iwamoto (Kansai University Graduate School), Shigenori Tanaka, Takeshi Naruo, Haruka Nishikoba, Yuhi Kameda (Kansai University)

A-3. Research on Visualization of Golf Green Undulation
Shigenori Tanaka, Takeshi Naruo (Kansai University), Ryohei Matsumura, Koyuru Tanaka (Kansai University Graduate School), Yuta Sasahara (Kansai University)

A-4. Research and Development of Swing Analysis Applications
Takeshi Naruo, Shigenori Tanaka, Yuhei Yamamoto (Kansai University), Wenyuan Jiang (Osaka Sangyo University), Akira Yokomichi, Norio Fujimoto, Toshihiro Akagi and Shingo Hakamata (People Software Corporation)

A-5. Development of an application to measure decision-making skill using Virtual Reality in basketball
Yasufumi Ohyama (Graduate School of Computer Science and System Engineering, Okayama Prefectural University; National Institute of Technology, Sasebo College), Kazushige Oshita (Okayama Prefectural University), Yuji Teshima and Reo Nakamura (National Institute of Technology, Sasebo College)

A-6. Augmentation and Prediction of Health Status Data Using Machine Learning
Toshiki Isogai and Katsunori Oyama (Nihon University)

A-7. Experimental investigation of the validity of muscle quality indicator based on the analysis of resistive components using bioelectrical impedance analysis
Kazushige Oshita (Okayama Prefectural University), Akihisa Hikita and Ryota Myotsuzono (Kyushu Kyoritsu University)

A-8. Situation Recognition of Animal Approaching Using Multimodal IoT Camera Systems
Ryo Tochimoto (Nihon University Graduate School) and Katsunori Oyama (Nihon University)

B-1. Development of Remote Monitoring System for Number of People Getting off Buses Using Video Camera Images
Yuya Matsui and Masaya Nakahara (Osaka Electro-Communication University)

B-2. Fundamental Research on Supporting Report of Inspection Results from Road Surface Deterioration Images Using ChatGPT
Kaito Andachi, Masaya Nakahara (Osaka Electro-Communication University), Yoshinori Tsukada (Reitaku University) and Yoshimasa Umehara (Setsunan University)

B-3. Fundamental Research on Method for Generating Real Scale 3D Shape Reconstruction Results Using Exif Images
Shota Yamashita, Masaya Nakahara (Osaka Electro-Communication University), Yoshinori Tsukada (Reitaku University) and Yoshimasa Umehara (Setsunan University)

B-4. Fundamental Study on Automated Counting Method of Turning Movement Counts at Intersections Using Video Images
Ryo Sumiyoshi (Graduate School of Engineering and Design, Hosei University), Ryuichi Imai (Hosei University), Yuhei Yamamoto (Kansai University), Masaya Nakahara (Osaka Electro-Communication University), Daisuke Kamiya (University of the Ryukyus) and Wenyuan Jiang (Osaka Sangyo University)

B-5. Prototype of Digital Twin Environment Using 3D Spatial Information Infrastructure Based on Spatial IDs
Kenji Nakamura (Osaka University of Economics), Yoshimasa Umehara (Setsunan University), Masaya Nakahara (Osaka Electro-Communication University) and Ryuichi Imai (Hosei University)

B-6. Application of machine learning to single-cell RNA sequencing provides the effective drugs against drug-tolerant persister cells in colorectal cancer
Yosui Nojima (Osaka University), Ryoji Yao (Japanese Foundation for Cancer Research) and Takashi Suzuki (Osaka University)

B-7. Dynamic Control of the Optimal Product Prices Using Machine Learning
Kazuhiro Miyatsu (Senshu University)

B-8. Empowering Educators: Assessing and Advancing Minecraft Education Edition Resources
Yan Zhu and Ryousuke Furukado (Nishinippon Institute of Technology)

B-9. What motivates evacuees to leave their cars during a tsunami evacuation?
Toshiya Arakawa (Nippon Institute of Technology)

Digital Life

Journal of Digital Life 2024, A-1

Conference Proceedings

Discussion on Detection of Dangerous Objects on the Road Surface Leading to Motorcycle Accidents Using a 360-degree Camera

Haruka Inoue [1,*] and Yuma Nakasuji [1]

[1] Faculty of Information Technology and Social Sciences, Osaka University of Economics, 2-2-8 Higashiyodogawa-ku, Osaka 533-0004, Japan

1. Introduction

In recent years, the number of fatalities in traffic accidents involving motorcyclists has remained almost unchanged, and although the Metropolitan Police Department has been conducting motorcycle safety classes, the number of fatalities increased for all ages in 2023. In addition, single-vehicle accidents accounted for 37.2% of all accidents by accident type in the past five years from 2018 to 2022 (Tokyo Metropolitan Police Department, 2024). Therefore, riders are required to follow the traffic rules and instantly predict danger. Themes of existing studies include development of an overturn prevention device for motorcycles using the gyro effect (Senoo et al., 2017) as well as detection of dangerous objects that may cause a motorcycle to overturn using deep learning. The former requires downsizing of the device, while the latter requires verification under different conditions. In this study, we apply a method to detect dangerous objects on the road surface from video images using YOLO to two types of 360-degree cameras (hereinafter referred to as "Dangerous object detection method") and verify that this method is versatile. In this study, as in existing studies (Inoue et al., 2023), fallen leaves, gravel, manholes, bumps, and wet road surfaces are considered as dangerous objects on the road surface.

2. Method

Figure 1 shows the process flow of the dangerous object detection method. The dangerous object detection method consists of a learning function and an estimation function. The input data for the learning function is the learning data, and the output data is the dangerous object detection model. The input data for the estimation function are video images taken by the 360-degree camera while riding a motorcycle, and the output data are the results of dangerous object detection. The learning function builds up a learning model to detect dangerous objects on the road surface that may cause a motorcycle to overturn. Specifically, as shown in Figure 2, the model to detect fallen leaves, gravel, manholes, bumps, and wet road surface as dangerous objects from video images (hereinafter referred to as "dangerous object detection model") by annotating dangerous objects on the road surface and learning them using YOLOv5. The estimation function is used to detect dangerous objects on the road surface from video images captured by the 360-degree camera. In the image generation process, the THETA+ application is used to convert the display format to flat, and crop to 1.91:1, and cut out the video image at 3 fps. The dangerous object detection process is used to detect dangerous objects on the road surface using the dangerous object detection model built up by the learning function.

3. Demonstration Experiment

In this experiment, we confirm the versatility of the proposed method by applying it to the video images captured using 360-degree cameras by a man in his 20s riding a motorcycle in an urban area in Wakayama Prefecture. First, two types of 360-degree cameras, THETA SC and THETA V both of which are products of RICHO, are installed at the front part the motorcycle (Figure 3). The ISO sensitivity differs between the two, with THETA SC ranging from

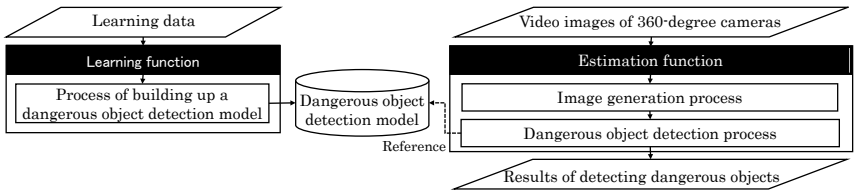


Fig. 1. Process Flow

[Legend] Red frame: fallen leaves, yellow frame: bumps

Fig. 2. Example of annotation

Published: 16 September 2024
* Correspondence: h.inoue@osaka-ue.ac.jp

100 to 1,600 and with THETA V from 64 to 6,400. The video resolution is 1,920 × 960 for both cameras. Next, in this study, since one 360-degree camera is installed at a time at the front part of one motorcycle in this study, the rider travels on the same section of road for multiple times at the same speed as much as possible. The results of detecting dangerous objects detected by applying the dangerous object detection method to the video images are compared with the manually generated correct-answer data, to make evaluation based on precision, recall, and F value. The learning model was built up using different images from the data used for the evaluation. For the learning data, the data taken by THETA SC on March 20, 21, July 10, and 12, 2024 were used. 3,103 images were used for the dangerous object detection model. On the other hand, the evaluation data were those taken by THETA SC and THETA V on March 22 and July 11, 2024.

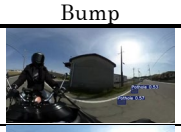
Table 1 shows the experimental results and Table 2 shows an example of the results of detecting dangerous objects. The overall F-measure is 0.897 for THETA SC and 0.872 for THETA V respectively, indicating that the dangerous object detection method is capable of detecting dangerous objects on the road surface correctly on the whole. Comparing the accuracy by camera type, the difference between THETA SC and THETA V was 0.015, which is not much difference. However, there were cases where only one of the 360-degree cameras was able to detect dangerous objects, even when the images were taken at the same point. Besides, comparing the accuracy by dangerous object type, the accuracy was lower for bumps and wet road surfaces than for fallen leaves. Focusing on the result of detecting dangerous objects, there was a tendency of omission of detection for small bumps or wet road surfaces covered with shadows. Therefore, we will increase the number of learning data under various environments and change the version of YOLO in order to advance the system.

Fig. 3.
Experimental view

Table 1. Experimental results

Dangerous objects		Fallen leave	Gravel	Manhole	Bump	Wet road surface	Total
THETA SC	Total number	20	23	11	47	11	112
	Number of determination cases	19	19	8	37	8	91
	Number of correct answers	19	19	8	37	8	91
	Precision	1.000	1.000	1.000	1.000	1.000	1.000
	Recall	0.950	0.826	0.727	0.787	0.727	0.813
	F-measure	0.974	0.905	0.842	0.881	0.842	0.897
THETA V	Total number	16	18	8	44	11	97
	Number of determination cases	15	13	7	33	7	75
	Number of correct answers	15	13	7	33	7	75
	Precision	1.000	1.000	1.000	1.000	1.000	1.000
	Recall	0.938	0.722	0.875	0.750	0.636	0.773
	F-measure	0.968	0.839	0.933	0.857	0.778	0.872

Table 2. Example of the results of detecting dangerous objects

Dangerous objects	Fallen leave	Gravel	Manhole	Bump	Wet road surface
THETA SC					
THETA V					

4. Conclusion

In this study, we applied the dangerous object detection method to the video images captured by 360-degree cameras and confirmed that the method is versatile enough to detect dangerous objects on the road surface even with different cameras or under different weather conditions and such through the demonstration experiment. In the future, we aim to further improve the accuracy by resolving the issues identified in this experiment.

References

Inoue, H., et al. (2023). Research for detecting dangerous objects leading to overturning of motorcycles using deep learning. *Proceedings of the 85th National Convention of IPSJ*, *85*(1), 965–966.

Senoo, D., et al. (2017). Development of motor-and-bicycle anti roll-down system. In *Proceedings of the Transportation and Logistics Conference* (Vol. 26). The Japan Society of Mechanical Engineers. https://doi.org/10.1299/jsmetld.2017.26.1104

Tokyo Metropolitan Police Department. (2024). Motorcycle traffic fatality statistics (through 2023). https://www.keishicho.metro.tokyo.lg.jp/kotsu/jikoboshi/nirinsha/2rin_jiko.html

Digital Life

Journal of Digital Life 2024, A-2

Conference Proceedings

Fundamental Research on Class Judgment of Preschooler in Early Childhood Education

Tatsuma Iwamoto [1], Shigenori Tanaka [2], Takeshi Naruo [3], Haruka Nishikoba [2] and Yuhi Kameda [2]

[1] Graduate School of Informatics, Kansai University Graduate School, 2-1-1 Ryozenji-cho, Takatsuki-shi, Osaka 569-1095, Japan
[2] Faculty of Informatics, Kansai University, 2-1-1 Ryozenji-cho, Takatsuki-shi, Osaka 569-1095, Japan
[3] Organization for Research and Development of Innovative Science and Technology, Kansai University, 3-3-35 Yamate-cho, Suita-shi, Osaka 564-8680, Japan

1. Introduction

In Japan, while the number of children using daycare facilities is decreasing due to the declining birthrate and aging population, the number of daycare facilities is increasing. However, the number of employees engaged in early childhood education has been decreasing year by year. The number of employees engaged in early childhood education is also decreasing. Consequently, reducing the workload of these employees has become an issue. Against this backdrop, the number of workers for each age group of children has been reviewed since the policy by the Administration for Children and Families in 2024, but the number of children each worker can take care of is still large, and there is a need to improve the "Quality of education" even further. Therefore, the use of ICT in early childhood education is attracting attention with the aim of reducing the workload. Currently, the use of ICT in early childhood education is limited to the education of children in classes using tablets and the support of paperless office work, and the use of ICT for work in which workers actually come into contact with children (watching over children) has not yet been implemented. Some facilities provide live video distribution of the facility, but this requires new work to check the video images and grasp the children, so there remains the issue of reducing the workload. In this study, we develop the basic technologies to support the monitoring of children by detecting and tracking children, and determining the class they belong to, using camera images installed in early childhood education facilities and AI technology. These methods will contribute to reducing the workload of those involved in early childhood education and to promoting the use of ICT in early childhood education in the future.

2. Proposed method

The processing flow of the proposed methods is shown in Fig. 1. In this study, we use video images captured by five cameras installed in preschool (Fig.2). Using the images from camera (1) to camera (4) and YOLOv7, we construct

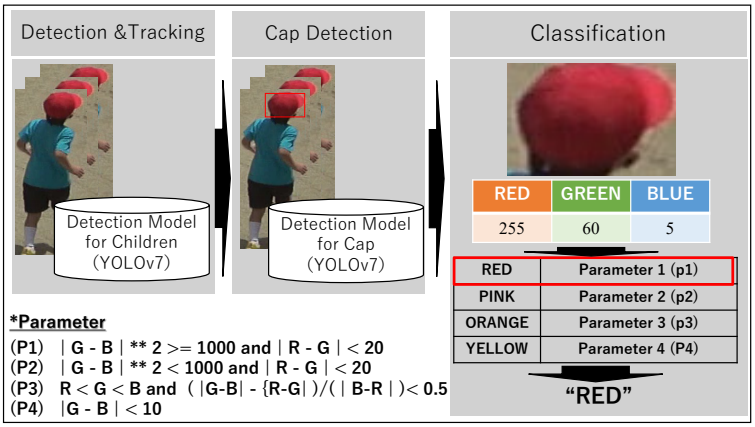


Fig. 1. Process flow of developed system

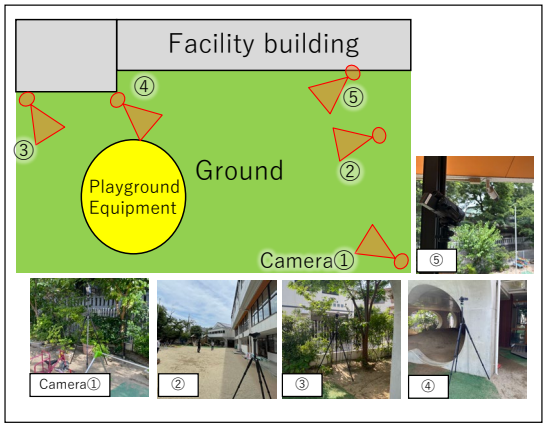


Fig. 2. Filming scenery camera setup

Published: 16 September 2024
* Correspondence: k153684@kansai-u.ac.jp

Table 1. Accuracy of detecting preschoolers present on the ground.

Model	Actual number of children	Counted result	Successful counting	Erroneous counting (Non detection)	Erroneous counting (False detection)	Precision ratio	Recall ratio	F-measure
COCO	2580	2038	1489	1088	549	0.73	0.58	0.64
ORIZIN	2580	2950	2495	89	410	0.86	0.97	0.91

Fig. 2. Examples of results of detecting preschoolers present on the ground.

Table 2. Classification results by cap color.

Cap color	Actual number of detected caps	Number predicted to be this color	Successful Judgment	Precision ratio	Recall ratio	F-measure
Red	518	489	467	0.96	0.90	0.93
Pink	432	444	396	0.89	0.92	0.90
Orange	376	395	376	0.95	1.00	0.98
Yellow	331	328	328	1.00	0.99	1.00

models for detecting children and their caps. For tracking, we use ByteTrack to track the children and detect their caps from each tracked data. The pixel values of the detected caps are extracted, the pixel values of the detected caps are extracted, and the color ranges of the pre-defined classes are summed. Then the child is identified as belonging to a class based on a majority vote of these summed color values.

3. Demonstration experiment

In this experiment, we confirm the usefulness of the proposed method using the 5-minute video from camera (5) in Figure 1. In the detection of preschoolers, we compare the results of our original model (ORIZIN) with those of the existing model (COCO), and visually confirm the accuracy. For class judgment, the accuracy is assessed by comparing the results with those of visual inspection.

The detection accuracy of the original model (Table 1) was found to be better than that of the existing model. The existing model tended to detect false positives in areas where there were no preschoolers. The original model also showed some false positives, but it was found that the model detected the areas around the children (feet and hands) incorrectly (Fig.2). The exiting model failed to detect 1,088 items, while the original model failed to detect only 89 items. Next, we examined the results of class identification (Table 2) and found that the class could be determined with high accuracy. Although there were some scenes where the class was misjudged, we believe that this can be improved by considering the time series and conducting majority voting within the tracking data.

4. Conclusion

In this study, we developed a method to detect and track preschoolers from camera images installed at early childhood education sites and to determine their class. Future work will involve verifying the system's applicability to long videos by introducing a person identification method.

Digital Life

Journal of Digital Life 2024, A-3

Conference Proceedings

Research on Visualization of Golf Green Undulation

Shigenori Tanaka [1], Takeshi Naruo [2], Ryohei Matsumura [3], Koyuru Tanaka [3] and Yuta Sasahara [1]

[1] Faculty of Informatics, Kansai University, 2-1-1 Ryozenji-cho, Takatsuki-shi, Osaka 569-1095, Japan
[2] Organization for Research and Development of Innovative Science and Technology, Kansai University, 3-3-35 Yamate-cho, Suita-shi, Osaka 564-8680, Japan
[3] Graduate School of Informatics, Kansai University Graduate School, 2-1-1 Ryozenji-cho, Takatsuki-shi, Osaka 569-1095, Japan

1. Introduction

In Japan, the utilization of sports and technology is being promoted to create new spectator experience and added value in sports. In golf as well, differentiation competition regarding the "visualization of the sports" has become increasingly prevalent among broadcasters, including the visualization of swing path and golf ball trajectory. Against this background, there is a demand for technology that can visualize green undulation information and the distance between the ball to the hole in real-time while filming and commenting on a player with broadcast footage. This information is valuable not only for spectators but also for athletes to improve the competitive performance. In particular, putting is a crucial element, requiring an accurate grasp of green undulation and strategizing based on this information. However, due to the limitations of conventional method (Sakuma et al., 2005), visually recognizing the green undulation quickly and easily has been difficult. Furthermore, there are challenges in obtaining undulation information by comparing it with actual video. Therefore, this research explores a method to automatically visualize undulation using machine learning with point cloud data generated from video captured by smartphones. Moreover, it aims to create CG video of putting scene with undulation information. Through this research, we seek to provide new information in golf broadcasts, contributing to the improvement of athlete performance and the technological revolution in sports broadcasting.

2. Methods

The processing flow of the proposed method is shown in Fig.1. The proposed method consists of an undulation estimation function and a CG video creation function. The former takes a green video as input and outputs undulation information. The latter utilizes a putting scene video as input and generates a CG video as output.

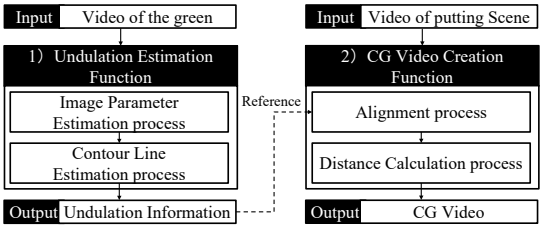


Fig.1. Processing flow

On the undulation estimation function, undulation information is obtained from video captured of the green. First, in the image parameter estimation process, point cloud data is generated from the green video using SfM (Schonberger et al., 2016), which reconstructs 3D data from multiple camera images. Next, in the contour line estimation process, unsupervised machine learning methods, k-means, is used to classify the height information of the point cloud data into multiple clusters (k). The optimal k value is selected through statistical analysis. Then, the boundaries of each cluster obtained with the selected k value are estimated as contour lines and added to the point cloud data as undulation information.

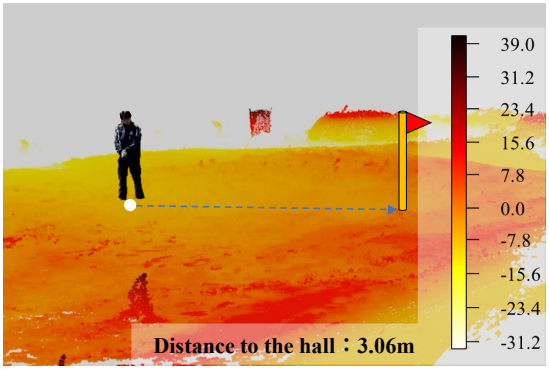


Fig.2. CG video

The CG video creation function calculates the distance from the ball to the hole and creates CG video of putting scene incorporating undulation information. First, in the alignment process, the point cloud data with undulation information is manually aligned with the putting video. Next, in the distance calculation process, the distance from the ball to the hole is calculated using position information from the point cloud data. Finally, CG video of putting scene incorporating undulation information is created as showed in Fig.2.

Published: 16 September 2024
* Correspondence: k119599@kansai-u.ac.jp

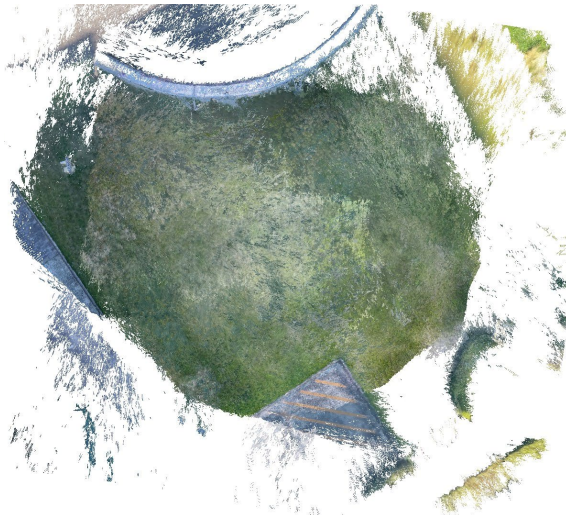


Fig.3. Result of SfM

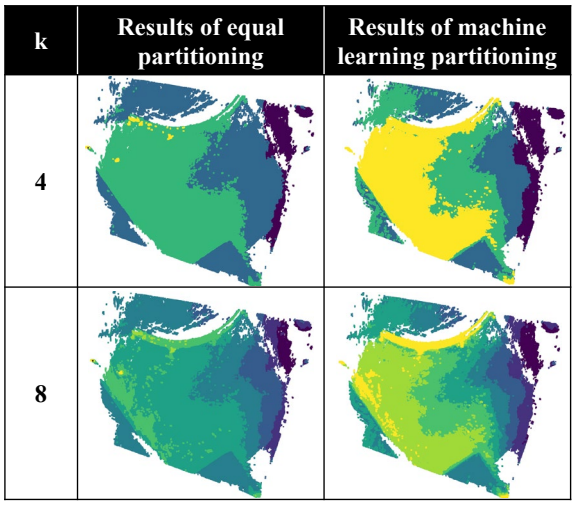


Fig.4. Results of Undulation Visualization

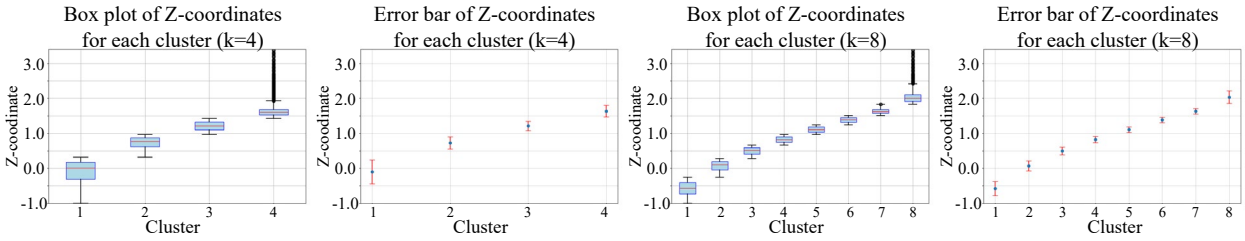


Fig.5. Statistical Analysis Results

3. Demonstration experiment

In this experiment, we use video of an environment simulating a golf course green as input data to generate CG video. First, regarding the undulation estimation function, we apply the proposed method to the point cloud model (Fig.3) created by the image parameter estimation process. We first compare the visualization results of the proposed method, referred to as "machine learning partitioning," with those of equal partitioning. The height of the target point cloud is divided equally based on the z-coordinate of the point cloud and color-coded to confirm its effectiveness. We use cases with 4 and 8 divisions as representative examples. Next, using the CG video creation function, we create CG video by overlaying it with video of putting. The experimental results are shown in Fig. 4. First, in the case of 4 divisions, the results of equal partitioning are susceptible to noise such as locally high areas, making it difficult to accurately identify detailed undulation on the green. However, machine learning partitioning enables more detailed confirmation of undulations on the green. Next, when the number of divisions is increased to 8, both methods make the green undulation easier to understand with the machine learning results appearing more robust against noise. Fig. 5 shows the box plot and error bars of the height-direction clusters in machine learning partitioning. Both the box plot and error bars indicate large error ranges for the minimum and maximum clusters, with the error range tending to decrease toward the middle range. The results demonstrate that while there is a slight reduction in accuracy in the fringe areas of the green, the division accuracy on the green itself is adequate. Regarding the CG video creation function, it was difficult to create real-time video due to the time required for manual overlay with camera footage. Therefore, by developing technology to automatically align point cloud data with the camera angle, real-time and easy CG video creation can be expected.

4. Conclusions

In this research, we proposed a method for visualizing golf green undulation. Through demonstration experiment, we confirmed that the proposed method enables visualization with a focus on the green. As future work, we will focus on enhancing both clarity and usability by addressing the challenges identified in this study.

References:

Sakuma, M., et al. (2005). The virtual image system of the golf green undulation for golf relays, *ITE Technical Report, 29*(9), 125–130.

Schonberger, L. J., et al. (2016). Structure-from-motion revisited, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 4104–4113. https://doi.org/10.1109/CVPR.2016.445

Digital Life

Journal of Digital Life 2024, A-4

Conference Proceedings

Research and Development of Swing Analysis Applications

Takeshi Naruo[1], Shigenori Tanaka[2], Yuhei Yamamoto[3], Wenyuan Jiang[4], Akira Yokomichi[5], Norio Fujimoto[5], Toshihiro Akagi[5] and Shingo Hakamata[5]

[1] Organization for Research and Development of Innovative Science and Technology, Kansai University, 3-3-35 Yamate-cho, Suita-shi, Osaka 564-8680, Japan
[2] Faculty of Informatics, Kansai University, 2-1-1 Ryozenji-cho, Takatsuki-shi, Osaka 569-1095, Japan
[3] Faculty of Environmental and Urban Engineering, Kansai University, 3-3-35 Yamate-cho, Suita-shi, Osaka 564-8680, Japan
[4] Faculty of Engineering, Osaka Sangyo University, 3-1-1 Nakagaito, Daito-shi, Osaka 574-8530, Japan
[5] People Software Corporation, 1-15-3, Achi, Kurashiki-shi, Okayama 710-0055, Japan

1. Introduction

In recent years, the rapid development of sports science and technology has significantly advanced techniques for enhancing athlete performance and preventing injuries. Consequently, the Ministry of Education, Culture, Sports, Science and Technology's 3rd Basic Sports Plan emphasizes the creation of innovative value in sports through the application of digital technology, positioning it as a key policy. Traditionally, high-performance sports motion analysis was mainly conducted at specialized research institutions and advanced training facilities, requiring expensive equipment. However, with the widespread use and improved performance of smartphones, advanced motion analysis is becoming accessible to general users. Therefore, in this research, we developed a system with advanced athlete motion analysis functions using Information technology. The developed system offers slow-motion playback, sequential photo output, and afterimage generation, making it versatile for analyzing various sports, especially swings in baseball and golf. Additionally, the developed system is not a dedicated application for expensive equipment but can be used with commonly available commercial smartphones, achieving affordability. This aims to promote the dissemination of sports analytics and contribute to improving the competitive skills of athletes across a wide range of levels.

2. Methods

As shown in Fig. 1, the proposed system consists of three main functions: motion replay function, afterimage output function, and sequential photo output function. In the motion replay function, users can freely adjust the playback speed to review the video. On commercially available smartphones, it is sometimes not possible to change the playback speed of recorded videos depending on the model. This makes it difficult to analyze fast movements in sports. To solve this problem, this function allows the playback speed to be adjusted anywhere from 0.1x to 8.0x, enabling the analysis of fast sports movements. The afterimage output function enables users to visualize the trajectory of movement by extracting the athlete's moving regions using interframe difference analysis of the input video. Then, by overlaying the regions from each frame, a single trajectory image is generated. This enables not only the assessment of the athlete's posture but also the evaluation of the quality of the entire sequence of movements. The sequential photo output function visualizes extracted continuous frame images for comparison. This allows users to conduct a more detailed analysis of a series of movements. In this function, users can select the pattern of image extraction. Four types of extraction patterns are provided according to the sport: equidistant extraction, front-biased extraction, center-biased extraction, and rear-biased extraction.

3. Results

To confirm the usefulness of each function, an actual baseball swing video was used. The video features a member of the Kansai University baseball team performing a swing. To evaluate the application, the recorded video was reviewed by the head coach, and feedback was collected regarding its instructional utility. The feedback results showed that many people evaluated this system as useful. The feedback on each function is summarized below.

Published: 16 September 2024
* Correspondence: tnaruo@kansai-u.ac.jp

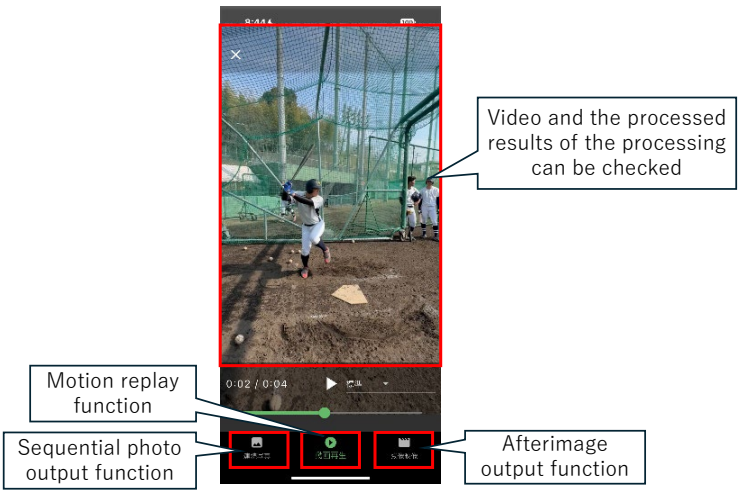


Fig. 1. Overall image of system

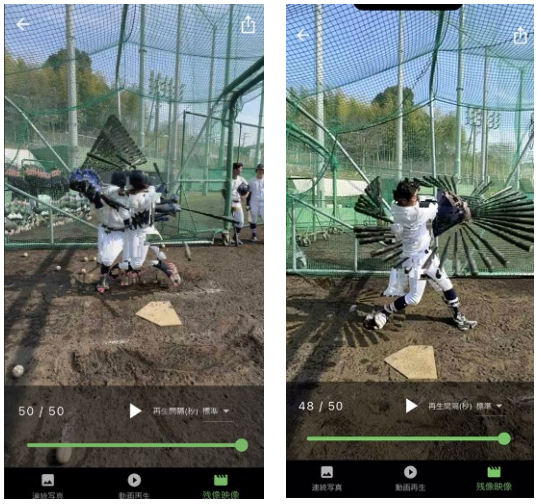


Fig. 2. Samples of generated afterimage

Fig. 3. Samples of the sequential photo output

4. Discussion

For the motion replay function, the ability to play a 240fps swing video at 8.0x speed, making it possible to view it at normal speed, allows users to easily review both slow-motion and normal-speed videos, enhancing swing analysis. Additionally, recording at 240fps enables the review of frames that cannot be seen at 60fps, allowing for more detailed analysis. This function was evaluated as very useful for coaching. The afterimage output function allows users to visually understand a problem with their swing trajectory (Fig. 2) by checking the swing path. This function was particularly effective for analyzing the synchronization between the head's trajectory and the body's rotation. It was evaluated as extremely useful for explaining and promoting understanding among athletes. In the sequential photo output function, as shown in sample A of Figure 3, baseball swings can be visualized through image extraction, allowing for point-by-point verification and thus enabling more detailed movement analysis. Additionally, as shown in sample B of Figure 3, it was found to be highly effective for analyzing the alignment of jumps in cheerleading. Overall, the developed application was suggested to have a high potential for contributing to the improvement of swing techniques. It was particularly evaluated as an effective tool for visually understanding and improving one's swing problem even without professional coaching. However, a challenge was noted regarding the generation speed of the afterimage output function, which takes about 30 seconds. This problem will be addressed and improved in the future.

5. Conclusions

The swing analysis application developed in this research combines high functionality with affordability, demonstrating the potential to promote the spread of motion analysis in amateur sports. The validation results confirmed its usefulness for athletes at various levels, suggesting that it could be a powerful tool, especially for sports enthusiasts seeking to improve their skills. With the future utilization of AI technology and application to other sports, further contributions to the dissemination of sports analytics and the enhancement of athletic performance are expected.

Digital Life

Journal of Digital Life 2024, A-5

Conference Proceedings

Development of an application to measure decision-making skill using Virtual Reality in basketball

Yasufumi Ohyama [1,2,*], Kazushige Oshita [3], Yuji Teshima [2] and Reo Nakamura [4]

[1] Graduate School of Computer Science and System Engineering, Okayama Prefectural University
[2] National Institute of Technology (KOSEN), Sasebo College
[3] Department of Human Information Engineering, Okayama Prefectural University
[4] Advanced Engineering Course, National Institute of Technology (KOSEN), Sasebo College

1. Introduction

'Decision-making skill' in invasive team sports is defined as the skill to perceive the situation during game play and to subsequently determine the optimal action. The 'decision-making test' measures a player's skill to make decisions in the game. A player with better situational decision-making skill has an advantage in the game. Decision-making tests usually use diagrams, photographs and videos to present the situation. However, such tests have the following problems: (1) the viewpoint is in the third person instead of the first person, and (2) (in the case of video) the viewing angle is different from the actual playing situation. Therefore, such tests do not fully reproduce actual play. This study developed a new decision-making test application to solve these problems, in the case of basketball. The application is a decision-making test using Virtual Reality (VR) and a head-mounted display (HMD), based on images captured by a 360° camera. Furthermore, a survey was conducted on the usability of the application in this study.

2. Development of a VR-based situation decision-making application

The VR-based decision-making application was created using Unity (Unity Technologies). The application consisted of a 'test creation mode' in which tests could be easily created on a PC by mouse operation and a 'testing mode' in which the created decision-making tests were performed on an HMD. Pre-recorded video images from the player's point of view by a 360° camera were used for the test. The situations of the images were recorded based on previous reports on decision-making tests. The video is paused at a moment that requires a situational decision-making during a play, then the question text and the answer choices are displayed. The participants select from the choices what they consider to be the appropriate next play.

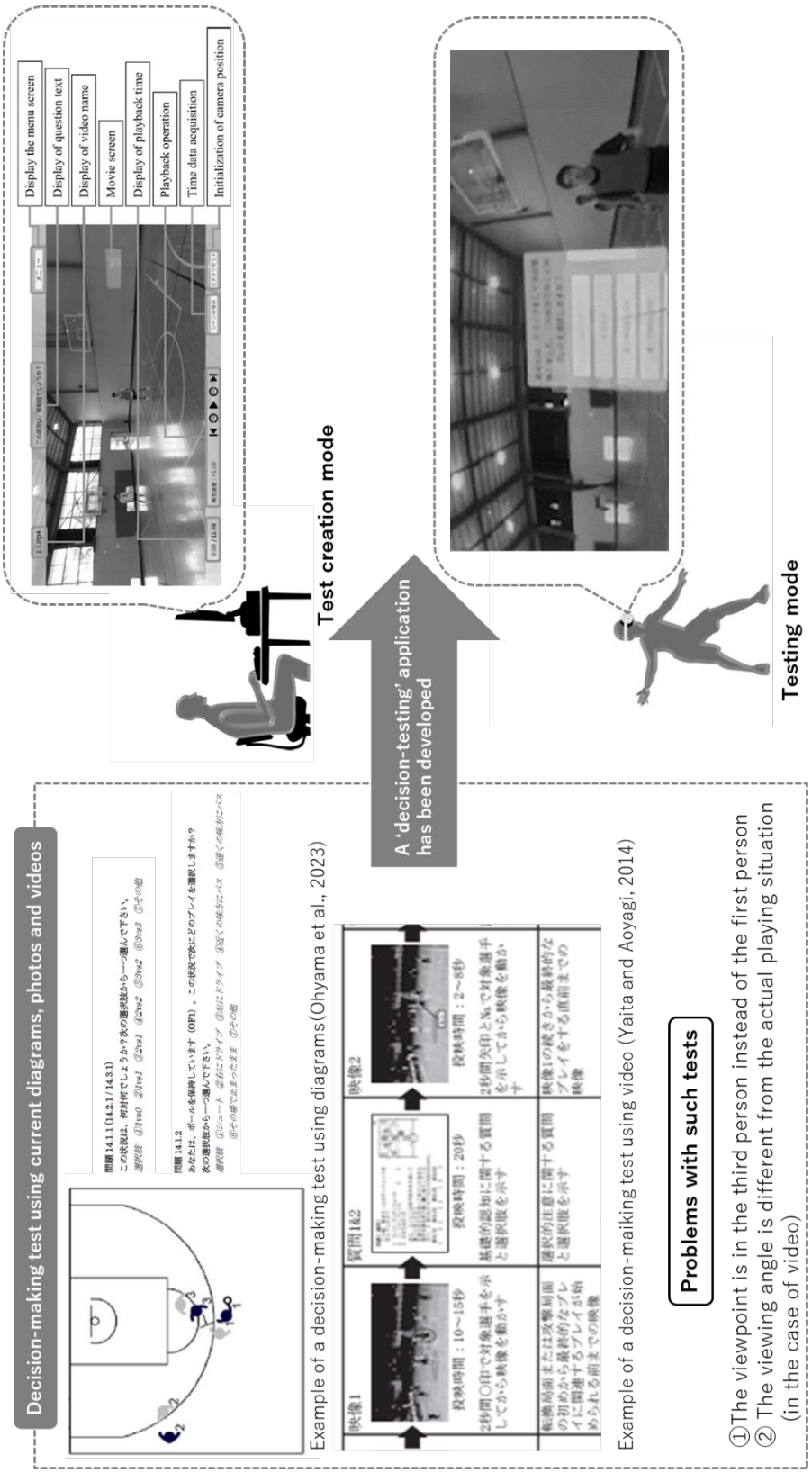
3. Survey of the usability of the VR-based test

Twenty-nine students from the technical college's basketball team were asked to experience a VR test. Subsequently, a questionnaire survey on the usability of the test was conducted. The results indicated that (1) approximately 97% of the participants answered that the VR test reproduced actual play, (2) approximately 93% of the participants answered that the VR test was smoothly performed, and (3) all participants answered that the VR test was useful in improving their decision-making skill.

4. Conclusions

The aim of this study was to develop a decision-making test using VR in basketball and to investigate the usability of the test among basketball players. This test application consisted of a 'test creation mode' and a 'testing mode'. The 'test creation mode' was designed to allow anyone to easily create tests on a PC using a mouse. The results of the usability survey of the VR test indicated that most participants found the test easy to use and well reproduced the actual play situation. Furthermore, all participants responded that the VR test was useful for improving their decision-making skills. These results suggest that the VR-based decision-making test application developed in this study can be easy to use and useful for improving decision-making skills.

Published: 16 September 2024
* Correspondence: yasufumi@sasebo.ac.jp



References

Ohyama, Y., Aoyagi, O., Yaita, A., Komure, I., Tagata, S., Nagamine, K., and Annoura, T. (2023) The relationship between decision-making ability in a basic confrontation with an opponent and various factors. *Research reports of National Institute of Technology, Sasebo College, 59*, 47–85.

Yaita, A., and Aoyagi, O. (2014) A factorial structure of decision-making ability during a fast-break in basketball games: Relationship between decision-making ability and game-styles of team, gender, athletic levels and positions. *Journal of training science for exercise and sport, 25*(2), 95–112.

Digital Life

Journal of Digital Life 2024, A-6

Conference Proceedings

Augmentation and Prediction of Health Status Data Using Machine Learning

Toshiki Isogai [1] and Katsunori Oyama [2]

[1] Graduate School of Computer Science, Nihon University
[2] Department of Computer Science, College of Engineering, Nihon University

1. Introduction

The increase in lifestyle-related diseases is becoming a major health concern. For example, the prevalence of diabetes, including those at risk, has surged by 1.2 times in the last five years, emphasizing the urgent need for effective countermeasures. These diseases can lead to severe complications such as Alzheimer's disease. However, current treatments typically focus on one condition at a time, often neglecting related issues. Therefore, preventive measures are crucial to reducing the incidence and risk of these diseases. Previous studies have demonstrated that blood tests can effectively predict cognitive function using the Mini-Mental State Examination (MMSE) as supervised data (Oyama & Sakatani, 2022). Recent research links cognitive function in the elderly with metabolic health including oral health (Karako et al., 2022). Nevertheless, developing accurate predictive models presents challenges due to high costs, small datasets, missing data, and limited availability of certain tests. This study aims to use machine learning to analyze and predict lifestyle-related diseases, their complications, and other health conditions using general blood test data and periodontal disease data.

2. Methods

In this study, we generate synthesized data using Generative Adversarial Networks (GAN) and analyze time series data related to lifestyle-related diseases.

2.1 Data Augmentation Using GAN

We used specific health checkup time series data from Habikino City (19,946 individuals) for data augmentation and created machine learning models including Random Forests. The dataset includes the presence or absence of 123 types of diseases for response variables. Then, GAN is employed to create synthesized data. To verify each machine learning model's performance and robustness against abnormal input values, we assess prediction accuracy with extreme values, such as Age (0 years and 100 years).

2.2 Time Series Analysis of Lifestyle-Related Diseases

First, we employed Random Forests to predict status changes of lifestyle-related disease. We classify the cases into three categories: (1) cases where there is a change in disease status from the initial examination, (2) cases where the disease is present in all examinations, and (3) cases where the disease is absent in all examinations. By classifying cases based on whether there is a change in the disease status or not, the prediction targets include not only disease recurrence and treatment but also cases where recurrence and treatment are repeated.

3. Results

The results of the robustness verification with abnormal input values showed that the models trained on real data had an increased MAE (Mean Absolute Error), while the models trained on synthesized data maintained a stable MAE. This indicates that models trained with synthesized data demonstrate robustness even when faced with abnormal data for important variables such as age and remaining teeth.

In the analysis of time series data related to lifestyle-related diseases, the accuracy was at least 64%. However, most items were classified with high accuracy. This suggests that it is possible to classify individuals based on changes in the presence or absence of diseases, enabling the identification of those at risk of developing a condition or those who may recover. However, the items that were fully classified had few cases and were mostly influenced by gender.

Published: 16 September 2024
* Correspondence: ceto23004@g.nihon-u.ac.jp

For prediction of status change in diabetes, as a representative lifestyle-related disease, the accuracy was 0.76, and HbA1C, which is related to diabetes, showed the highest feature importance (0.09). This indicates that similar analyses for other diseases could lead to the discovery of important features related to lifestyle-related diseases and other conditions. On the other hand, for prediction of status change in kidney failure, the accuracy was 0.99, and CREA (creatinine) showed the highest feature importance as 0.11. Creatinine is a substance filtered by the kidneys and excreted from the body, closely related to kidney function, which indicates that the feature importance of the model is related to each disease. Furthermore, among the items with high feature importance for other diseases, many diseases showed high feature importance for age, which suggests that the onset of these diseases is highly related to aging.

Table1. Top 10 feature importances for prediction of kidney failure and diabetes.

Kidney failure (Accuracy: 0.99, F1 score: 0.98)		Diabetes (Accuracy: 0.76, F1 score: 0.72)	
Feature	Importance	Feature	Importance
CREA	0.11	**HbA1C**	0.09
HEIGHT	0.06	GlC	0.07
UA	0.05	AGE	0.06
WEIGHT	0.05	WEIGHT	0.05
RBC	0.05	HEIGHT	0.05
HDL-C	0.05	LDL-C	0.05
AGE	0.05	TG	0.05
TG	0.05	CREA	0.05
LDL-C	0.05	RBC	0.05
HT	0.05	HDL-C	0.05

4. Discussions

The comparison of machine learning models using real data versus synthesized data has shown that synthesized data can be used for predictions similarly to real data. Furthermore, using synthesized data improved the model's robustness, demonstrating its robustness. This suggests a potential for training machine learning models with synthesized data. Additionally, in predicting lifestyle-related diseases, the accuracy of predicting changes in disease status was at least 63%, and nearly 90% for most diseases, indicating that the model can effectively distinguish between most diseases. However, for items with 100% classification accuracy, consideration must be given to factors such as diseases being gender-specific or having a small number of cases and limited samples of affected individuals, and thus further evaluation of these factors is necessary. One of the factors contributing to the increase in lifestyle-related diseases is the rise in the elderly population, which might be a cause according to the results of this experiment. Moreover, for other factors, the high importance of features related to conditions like renal failure and diabetes suggests that it is possible to identify key features associated with other diseases as well.

5. Conclusions

In this study, the use of GANs allowed for an increase in the size of the training data. While the real data was biased, with a concentration in the 50s to 70s age range, GAN-synthesized data created a more uniform distribution. However, despite this, the anticipated improvement in prediction accuracy was not as significant as expected. This suggests that further improvements are needed in both the GAN and DNN models. In robustness testing, the models trained with GAN-synthesized data demonstrated robustness, showing that they were capable of handling abnormal input values. The analysis of time series data related to lifestyle-related diseases indicated the potential for predicting changes in disease status from health checkup data. Additionally, it is anticipated that this approach will help in identifying features affecting each disease.

References

Karako, K., Chen, Y., Oyama, K., Hu, L., & Sakatani, K. (2022). Relationship between cognitive function, oral conditions, and systemic metabolic function in the elderly. *Advances in Experimental Medicine and Biology, 1438,* 27–31. https://doi.org/10.1007/978-3-030-65928-4_4

Oyama, K., & Sakatani, K. (2022). Machine learning-based assessment of cognitive impairment using time-resolved near-infrared spectroscopy and basic blood test. *Frontiers in Neurology*, *12*, 624063. https://doi.org/10.3389/fneur.2021.624063

Digital Life

Journal of Digital Life 2024, A-7

Conference Proceedings

Experimental study of the validity of muscle quality assessment based on the analysis of resistive components using bioelectrical impedance analysis

Kazushige Oshita[1]*, Akihisa Hikita[2] and Ryota Myotsuzono[2]

[1] Department of Human Information Engineering, Okayama Prefectural University
[2] Department of Sports Science, Kyushu Kyoritsu University

1. Introduction

Bioelectrical impedance analysis (BIA) is a simple, non-invasive technique for assessing body composition, including body fat and muscle mass, and is widely utilised for home-based monitoring. BIA is a method for estimating body composition based on electrical resistance (impedance, Z) by applying a weak alternating current to the body and analysing the characteristics of high electrical resistance in fat tissues containing a small amount of water with low electrical resistance in muscle tissues containing water. The assessment of muscle mass is crucial in several contexts, including the evaluation of sports performance, normal weight obesity and the diagnosis of sarcopenia in the elderly. However, recent studies on body composition have focused on the assessment of not only 'muscle mass', but also 'muscle quality.'

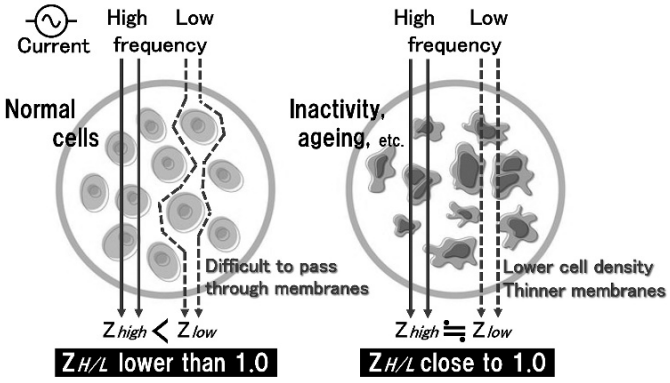


Fig. 1 Differences in current flow according to cell status

Skeletal muscle tissue is a collection of muscle cells, and each cell membrane (primarily composed of phospholipids) exhibits electrical resistance properties (reactance). The BIA method demonstrates that low-frequency currents (*f*Lc) cannot pass through the membrane, whereas high-frequency currents (*f*Hc) can pass through the membrane (Ellis, 2000) (Fig. 1, left). Therefore, the *f*Lc evaluates the extracellular water content, whereas the *f*Lc evaluates the intracellular and extracellular water contents. *f*Lc in muscles with a high number of cells create an extended conductor length owing to increased bypass (that is, higher Z), and vice versa in muscles with a low number of cells. Furthermore, cell deterioration (such as thinner cell membranes) associated with aging or inactivity allows *f*Lc to pass through membranes more easily (Fig. 1, right). Because Z of the *f*Lc (Z_{low}) is low in muscles with fewer or deteriorated cells, its difference with the Z value of the *f*Hc (Z_{high}) is presumed to be smaller (Fig. 1). Therefore, the ratio between Z_{high} and Z_{low} ($Z_{H/L}$) can be used to evaluate the state of muscle cells.

This study investigated whether the cellular state of muscles can be evaluated using $Z_{H/L}$. The frequencies corresponding to *f*Lc and *f*Hc in commercially available BIA body composition analysers were 5k and 250k Hz, respectively. In this study, $Z_{H/L}$ was calculated using these two frequency currents.

2. Experimental study on resistive components using BIA method

A BIA device with electrodes placed on both hands and feet (MC-780, Tanita, Japan) was used in this study. The value of Z is proportional to volume resistivity and conductor length, and inversely proportional to conductor cross-sectional area. The resistivity of the BIA method indicates the body composition (percentage of muscle and fat). The conductor length was calculated as the approximate distance between the electrodes based on the pre-inputted height. The body composition of 15 healthy male participants was evaluated the following two conditions: (1) a normal measurement

Published: 16 September 2024
* Correspondence: oshita@ss.oka-pu.ac.jp

with the limbs extended (NORMAL), and (2) a measurement with the armpits closed and the upper limbs flexed (FLEX) (Fig. 2). Z was measured from the current in the left half of the body. The results demonstrated that the upper arm circumference exhibited a significantly greater increase in the FLEX condition relative to the NORMAL condition. Additionally, both the Z_{high} and Z_{low} exhibited a significant decreased by approximately 10%, and the displayed body fat percentage (%BF) also decreased. While a decrease in $Z_{H/L}$ was also observed, it was only approximately 2% (Fig. 2).

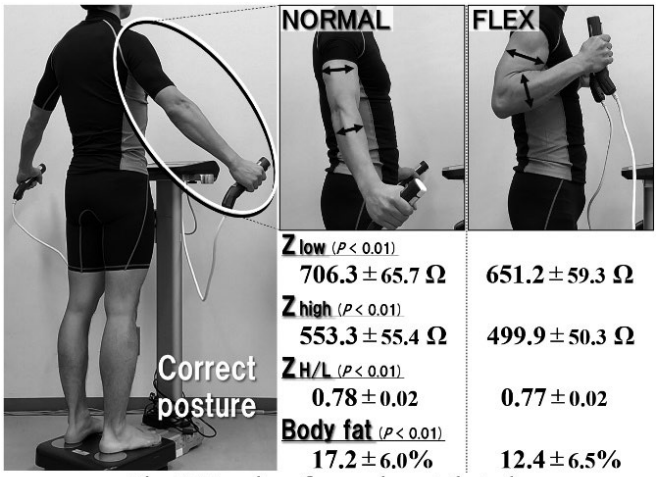


Fig. 2 Results of experimental study

Under FLEX conditions, the cross-sectional area increased because of the arm flexion, whereas the conductor length decreased owing to the formation of contact between the arm and torso or forearm and upper arm. Therefore, the Z and %BF values were lower under the FLEX condition than under the NORMAL condition. However, the number of cells remained unchanged when the arms were flexed (but the membrane exhibiting slight stretching and thinning owing to the compression of the cells via flexion). Consequently, a more pronounced decrease in Z was observed under the FLEX conditions, whereas the decrease in $Z_{H/L}$ was relatively minimal.

3. Comparison of body composition and $Z_{H/L}$ based on participant characteristics

A cross-sectional study was conducted to investigate the relationship between $Z_{H/L}$ and different physical activity level. The study examined 46 university students majoring in non-sports subjects (e.g., information technology and nutrition) and 93 students majoring in sports. The sports students were divided into three groups: 31 who had competed in university national championships in power events, such as track and field throwers (elite), 33 with a lower body mass index (BMI) (≤ 22 kg/m^2) (L-sports), and 29 with a higher BMI (≥ 23 kg/m^2) (H-sports). Body composition was assessed using the BIA method, in addition to the BMI, %BF, fat-free mass (FFM) and $Z_{H/L}$, which were then compared.

The elite group exhibited a significantly higher BMI, %BF, and FFM than the other groups (Table 1). Although the H-sports group displayed significantly higher BMI, %BF, and FFM than the L-sports group, $Z_{H/L}$ did not show a statistically significant difference. However, $Z_{H/L}$ was significantly higher in the L-sports group than in the non-sports group despite the lack of statistically significant differences in the BMI and FFM between these two groups.

Table 1 Comparison of body composition and $Z_{H/L}$ according to participant characteristics

	elite (n = 31)	H-sports (n = 33)	L-sports (n = 29)	non-sports (n = 46)	ANOVA
BMI (kg/m^2)	29.6 ± 5.0 *† ‡	25.5 ± 3.0 *†	21.5 ± 1.2	22.2 ± 3.3	$F = 39.88$, $P < 0.01$
% Body fat (%)	19.8 ± 5.6 *† ‡	16.8 ± 5.4 †	11.9 ± 3.7 *	16.6 ± 5.8	$F = 12.70$, $P < 0.01$
Fat-free mass (kg)	74.2 ± 8.6 *† ‡	62.3 ± 6.0 *†	54.8 ± 5.9	54.2 ± 6.9	$F = 60.68$, $P < 0.01$
$Z_{H/L}$	0.73 ± 0.02 *† ‡	0.75 ± 0.01 *	0.75 ± 0.02 *	0.77 ± 0.02	$F = 20.02$, $P < 0.01$

Values are means ± S.D.
* vs. non-sports, † vs. L-sports, ‡ vs. H-sports, $P < 0.05$ (Holm's method)

4. Conclusions

This study indicates that the ratio of Z between *f*Lc and *f*hc using the BIA method ($Z_{H/L}$) facilitates assessment at the mscule cell status, which cannot be assessed using only muscle mass. In particular, (1) $Z_{H/L}$ was higher in students majoring in sports than those majoring in non-sports subjects, despite no difference in physique or muscle mass between both non-sports and L-sports groups, and (2) both $Z_{H/L}$ and muscle mass were higher in the elite group than in the H- and L-sports groups. These results are interesting, as $Z_{H/L}$ may be an indicator of 'muscle quality.' Although a cross-sectional design was adopted in this study, future longitudinal studies are essential to elucidate the causal relationship between $Z_{H/L}$ and muscle quality.

Acknowledgement

This research was supported by The Okayama Foundation for Science and Technology and JSPS KAKENHI Grant Number JP24K09636.

References

Ellis, K. J. (2000). Human body composition: in vivo methods. *Physiological reviews, 80* (2), 649–680.

Digital Life

Journal of Digital Life 2024, A-8

Conference Proceedings

Situation Recognition of Animal Approaching Using Multimodal IoT Camera Systems

Ryo Tochimoto [1], Katsunori Oyama [2]

[1] Department of Computer Science, Faculty of Engineering, Nihon University Graduate School
[2] Department of Computer Science, Nihon University, Koriyama, Japan

1. Introduction

Crop damage by wild animals has become a serious social problem, and it is important to observe the behavior of pests as a countermeasure against pests (Ministry of Agriculture, Forestry and Fisheries, 2022). However, it is difficult to capture wild animals within the angle of view of a camera, and even if an infrared sensor is used, false detection and battery depletion are problems. In this study, we propose an efficient situation recognition method by constructing an IoT camera system that combines a Raspberry Pi, an infrared sensor with analog output, and an audio microphone, and using machine learning to classify multimodal data obtained from the surrounding environment.

2. Situation Awareness Model for Animal Approach

In the field of situation awareness, the definition of situation has evolved, and this paper discusses situation as not only as classification results from machine learning models but also the actions and environmental contexts with those results (Chang et al., 2009). Proper understanding became an emerging issue for analyzing the behavior of wild animals (Tochimoto et al., 2023) In this approach, our IoT camera system is constructed using a combination of Raspberry Pi, infrared sensors, and audio sensors, interpreting complex information as situational data. The system integrates data from multiple sensors and classification results to assess "animal approaching".

3. IoT Camera System

The IoT camera previously developed in this project is based on the Raspberry Pi Zero2 and integrates an infrared sensor, infrared camera, Raspberry Pi camera, and audio microphone (Tochimoto et al., 2023). In this study, the dedicated hardware, equipped with various sensors, is re-designed to be power-efficient and compact (Fig. 1). A custom case was designed using a 3D printer, enhancing the system's waterproof capabilities. The battery and sensors are housed in a small, integrated design.

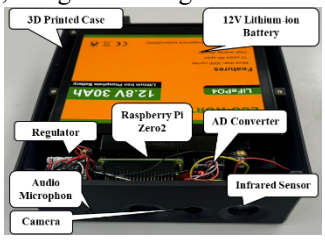


Fig. 1. Internal Structure

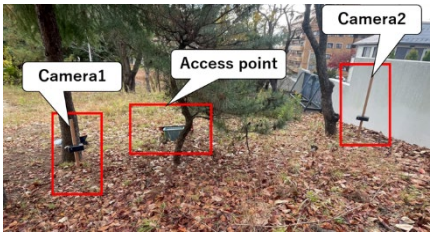


Fig. 2. Experimental Environment

The infrared sensor used is a pyroelectric infrared sensor PaPIRs (manufactured by Panasonic, 12m long-range detection type), capable of collecting time-series data from the surrounding environment, which traditional sensors could not capture. The system also includes an omnidirectional microphone to collect environmental sounds, such as rain, wind, and animal calls, which are classified through frequency analysis. The system employs multithreading to facilitate the simultaneous operation of infrared data collection, cameras, and microphones. The infrared sensor readings are taken every 0.1 seconds, and data collection begins when a reaction is detected. Additionally, the system uses Raspberry Pi's cron process to enable automatic sensing upon power-on.

Published: 16 September 2024
* Correspondence: cery23018@g.nihon-u.ac.jp

4. Experiments

The IoT camera continuously monitors infrared waveforms and captures images, videos, and audio when a threshold is exceeded. The collected data is saved as CSV files every hour and can be retrieved remotely via a 4G connection. Experiments were conducted at the site of the Nihon University Training Center, with the camera systems connected via Wi-Fi and positioned around a solar panel-equipped power source and a 4G router (Fig. 2).

Two test sites were used: Nihon University Training Center (April 2023) and Katsurao Village, Fukushima Prefecture (May 2024). Two cameras were installed at each site, and data was collected within the designated range. Fourier transform (FFT) was applied to the collected infrared and audio waveform data to extract features. This process calculated the sum and variance of the data and created labeled analysis data, distinguishing between successful detection (Approaching) and failed detection (No Approaching). For preprocessing audio data, the first 0.3 seconds after recording start were cut to remove noise.

5. Results

At Nihon University Training Center, 267 data points were collected, with 13 instances of animal presence and 254 instances of absence. In Katsurao Village, 366 data points were collected, with 73 instances of animal presence and 293 instances of absence. The captured images included humans, cats, and raccoons, with false detections also occurring due to wind and other factors.

A model was constructed using infrared and audio waveform features to classify the presence or absence of animals. Data collected at the Nihon University Training Center was used as training data, and data from Katsurao Village was used for validation. Table 1 shows the results of classification using each model. Logistic regression and random forest models accurately detected the presence of animals, while SVM struggled to detect the absence of animals.

Table 1. Learning Results of Random Forest (RF), Logistic Regression (LR), and Support Vector Machine (SVM).

Actual \ Predicted	RF (Sensitivity:0.58, Specificity:0.74)		LR (Sensitivity:0.29, Specificity:0.84)		SVM (Sensitivity:0.93, Specificity:0.05)	
	Negative	Positive	Negative	Positive	Negative	Positive
Negative	99	34	112	22	0	133
Positive	31	42	52	21	0	73

Comparing the feature importances, both the random forest and logistic regression models utilized a balanced mix of infrared and audio data, whereas the SVM model leaned heavily on infrared data. This balance in the models helped in accurately detecting animal approaches and distinguishing them from environmental noise. Analysis of high-frequency bands was crucial for accurately detecting the presence of animals.

Infrared sensors detected changes in thermal radiation due to animal body temperature and movement, particularly identifying high-frequency components as key indicators of movement. Similarly, high-frequency components in audio waveforms were effective in identifying sounds related to animal movements, filtering out low-frequency noise from environmental factors. These findings highlight the effectiveness of the IoT camera system and indicate areas for further improvement.

6. Conclusions

Using the multimodal IoT camera, the experimental result indicated the potential for wildlife management studies. The system demonstrated its effectiveness by continuously collecting infrared and audio waveform data in outdoor environments. Analysis results confirmed the ability to effectively distinguish animal proximity. Future challenges include enhancing the sample data and extending the system's operating time.

References

Chang, C. K., Jiang, H., Ming, H., & Oyama, K. (2009). Situ: A situation-theoretic approach to context-aware service evolution. *IEEE Transactions on Services Computing, 2*(3), 261–275. doi: 10.1109/TSC.2009.21

Ministry of Agriculture, Forestry and Fisheries. (2022, November 10). *Report on Crop Damage Caused by Wildlife Nationwide (FY2020).*

Tochimoto, R., Oyama, K., & Ming, H. (2023). Development of an IoT camera system for situation recognition of approaching animals. *IEEE International Conference on Software Services Engineering (SSE)*, 1–6. doi: 10.1109/SSE60056.2023.00047

Digital Life

Journal of Digital Life 2024, B-1

Conference Proceedings

Development of Remote Monitoring System for Number of People Getting off Buses Using Video Camera Images

Yuya Matsui [1]and Masaya Nakahara [1]

[1] Faculty of Information Science and Arts, Osaka Electro-Communication University, 1130-70 Kiyotaki, Shijonawate-shi, Osaka 575-0063, Japan

1. Introduction

In Japan, the number of users of buses, a means of public transportation, has been decreasing due to the nationwide increase in the use of automobiles in addition to the decline in population (Ministry of Land, Infrastructure, Transport and Tourism, 2021). Against this backdrop, many bus companies have been improving the efficiency of their bus operations by surveying the number of passengers and reviewing their route plans. In addition, universities and other facilities that operate buses on behalf of bus companies purchase data on the number of passengers on buses from the bus companies, which can be very costly on an ongoing basis. For this reason, they request bus operations based on the number of bus passengers counted manually. However, the long hours of outdoor monitoring required for manual surveys may cause health hazards to the surveyors due to hot and cold temperatures, and human error may occur. To prevent these problems, surveys are conducted by multiple persons, but this incurs huge labor costs. Therefore, there is a need for a technology to automatically count the number of people getting off the bus.

Existing studies have proposed methods such as counting using a bus-mounted drive recorder (Nakashima et al., 2018) and a method for counting by installing area-measuring sensors inside the bus (Yamada et al., 2019). In these methods, passengers passing near the farebox are counted mechanically using drive recorders and sensors installed around the bus's drop-off entrance. However, when buses are crowded, occlusion among passengers may result in false counts due to omissions or tracking errors. In addition, when facilities or groups other than bus companies survey the number of passengers who get off the bus, they must borrow drive recorders and sensors managed by the bus companies, which incurs continuous financial costs.

2. Methods

Based on the issues with existing methods, this study develops a system that can recognize and count the number of passengers getting off a bus using deep learning from video taken around a bus stop, and remotely monitor the number of people getting off via a web page. Specifically, we first construct a model that detects buses, bus exits, and people in the video using deep learning. Next, using the model, the system detects buses and their exits in the video. The bus with the largest rectangular area among the detected buses is estimated to be the bus stopped closest to the video camera and is counted. Based on the color of the bus body, the bus company is estimated based on the color threshold set in advance. Next, using the constructed model, a person is detected in the video image, and the same person is linked and tracked from frame to frame based on the person's rectangle and skeletal structure. Based on the location of the person and the exit of the bus, the system then estimates whether the target person is a passenger who has alighted from the bus, and if so, records the number of passengers. Finally, if the person is not recorded for a certain period of time, the system assumes that all passengers have alighted from the bus and displays the results of counting the bus company and alighted passengers on a web page.

3. Demonstration experiment

In this experiment, the proposed system is used to count the number of people getting off a bus based on video images taken around a bus stop for several hours, and the results are compared with those obtained by visual counting. We

Published: 16 September 2024
* Correspondence: nakahara@oecu.jp

n台目	到着時刻	バス会社	降車人数
1	2024年7月8日 - 21時6分57秒	京阪バス	40
2	2024年7月8日 - 21時13分48秒	近鉄バス	51
3	2024年7月8日 - 21時41分33秒	京阪バス	12

Fig. 1. Web page displaying counting results

Table 1. Counting results

No	Correct Bus company	Bus estimation results	Number of correct answers	Number of people counted	Number of correct answers within count	Precision	Recall	F-values
1	Keihan Bus	Keihan Bus	41	40	40	1.000	0.976	0.988
2	Kintetsu Bus	Kintetsu Bus	52	51	51	1.000	0.981	0.990
3	Keihan Bus	Keihan Bus	12	12	12	1.000	1.000	1.000

Table 2. Hearing results

Evaluation of current functionality	・groundbreaking. I want to actually install and use it. ・Useful as an indicator when requesting service from bus companies.
Opinions on visualization methods	・If there are more than 50 people on the bus, I want highlights. ・If you arrive more than 5 minutes late, I want highlights.
Opinions on aggregation method	・I want to download the aggregated data. ・I also want to know the number of passengers. ・I want to know the bus system number (direct bus, etc.). ・Simultaneous counting of multiple units.

will then confirm the usefulness of the proposed system through interviews with practitioners and organize issues to be addressed for its practical use.

The video images taken around the bus stops and the results of the analysis using the proposed system are shown in Fig. 1, and the results of the counting are shown in Table 1. The goodness-of-fit rate, recall rate, and F value are all greater than 0.988. This confirms the usefulness of the system. However, the reproducibility of No. 1 and No. 2 was 0.976 and 0.981, indicating an omission in counting. This was due to the fact that passengers got off the bus between the arrival of the bus and the start of the counting process. Therefore, it is necessary to estimate whether a bus is stopped or not using the movement of the entire bus so that it can be estimated more quickly. Table 2 shows the results of the interview survey. We also obtained requests for visualization and tabulation methods. We plan to improve the proposed system based on these opinions and requests.

4. Conclusion

In this study, we developed a system that can count the number of people getting off a bus from video images taken around a bus stop and monitor the number via a web page. Through demonstration experiments, we were able to obtain high counting accuracy and opinions from practitioners and confirmed the usefulness of the proposed system. Since issues for practical use of the system were also identified, we will improve the proposed system to increase its applicability to real-world situations.

References

Ministry of Land, Infrastructure, Transport and Tourism. (2024). Movement of People in Cities and its Ch anges: From the Aggregated Results of the *2021 National Urban Transportation Characteristics Survey*. https://www.mlit.go.jp/report/press/content/001711623.pdf.

Nakashima, H., Arai, I., and Fujikawa, K. (2018). Estimating Number of Passengers by Background Subtraction Using Images of Drive Recorder Inside Buses. *DICOMO2018*, 43-48.

Yamada, Y., Hiromori, A., Yamaguchi, H., and Higashino, T. (2019). Development of Bus Passenger Counter Using LiDAR Sensors. *IPSJ Journal*, *60*(3), 934-944.

Digital Life

Journal of Digital Life 2024, B-2

Conference Proceedings

Fundamental Research on Supporting Report of Inspection Results from Road Surface Deterioration Images Using ChatGPT

Kaito Andachi [1], Masaya Nakahara [1], Yoshinori Tsukada [2] and Yoshimasa Umehara [3]

[1] Faculty of Information Science and Arts, Osaka Electro-Communication University, 1130-70 Kiyotaki, Shijonawate-shi, Osaka 575-0063, Japan
[2] Faculty of Engineering, Reitaku University, 2-2-1 Hikarigaoka, Kashiwa-shi, Chiba 277-8686, Japan
[3] Faculty of Business Administration, Setsunan University, 17-8 Ikedanakamachi, Neyagawa-shi, Osaka 572-8508, Japan

1. Introduction

In Japan the total length of roads is approximately 128 km. These roads are repeatedly damaged by vehicle traffic. As road damage spreads, serious problems such as traffic accidents can occur. To prevent such problems from occurring, road maintenance workers are required to keep track of the condition of roads by conducting periodic inspections. However, it takes an enormous amount of time to inspect all roads. Furthermore, there is concern that the number of inspection companies will taper off due to the declining birthrate and aging population. Therefore, there is a need for technology to reduce manpower and improve efficiency.

In existing research (Imai et al., 2021), a method for evaluating road cracks using deep learning from video images acquired by a drive recorder was proposed. The method evaluates cracks in three stages and visualizes the evaluated areas. In addition, a method (Asakawa et al., 2012) has been proposed to estimate the condition of a road traveled by a vehicle based on the vehicle's reactions while traveling. In this method, the dynamic response of a vehicle moving at a constant speed is captured and evaluated by vertical acceleration. Although quantitative evaluation methods have been proposed in existing methods, they do not output textual information such as findings.

In this research, we propose a method to support the preparation of reports by outputting the findings of inspections from images taken of deteriorated road surfaces using ChatGPT.

2. Methods

This chapter describes the flow of the method to support the preparation of a report using ChatGPT. Input data are images of deteriorated road surfaces. The output data is a textual description of the findings on the road surface condition.

ChatGPT is trained with images of road surfaces and sentences of findings. Next, we input images of the road surface to the trained ChatGPT. This allows us to estimate the condition of the road surface and to generate findings and other information. The ChatGPT used in this method is a fine-tuned using GPTs that can be customized to recognize road surface conditions. During the fine-tuning process, 100 images were trained for each of four types of images: images with cracks, images without cracks, and images with clear and blurred white lines. The images used for training were visually checked for cracks and white lines and labeled. The prompts were set so that the output text did not mention the need for repair due to the shadows in the pictures or the condition of the road.

3. Demonstration experiment

In this experiment, we verify whether the proposed method can discriminate road surface conditions and output appropriate findings. Fifty images of each of the six types shown in Table 1 (a) were prepared for evaluation. The resolution of these images is 3024pix×4032pix. Table 1 (a) shows the percentage of correct responses for the output results, and Table 1 (b) shows the percentage of correct responses for each road condition.

Published: 16 September 2024
* Correspondence: nakahara@oecu.jp

Table 1. Percentage of correct answers

(a)

	Cracks present	No cracks
Clear white line	0.62	0.90
Unclear white line	0.90	0.70
No white line	0.96	0.96

(b)

Road condition	percentage of correct answers
Cracks present	0.87
No cracks	0.88
Clear white line	0.92
Unclear white line	0.98
No white line	0.98

Table 1 (a) shows that the correct response rate for most of the images was over 90%, while the correct response rate for images with cracks and clear white lines was about 60%, and that for images with no cracks and unclear white lines was about 70%. This is thought to be due to the fact that the cracks prevented normal discrimination. In addition, when we checked the images in which the white lines were incorrectly estimated to be indistinct, we found that they were judged to be broken white lines due to cracks on the boundaries of the white lines. On the other hand, the image that was incorrectly estimated to have a crack was judged to have a crack due to an indistinct white line. Figure 1 (a) shows that in the image with no white lines and no cracks, the shadows of tree branches and fallen leaves were judged to be cracks. In the case of the image with cracks, the cracks were judged to be white lines, as shown in Figure 1(b). Table 1 (b) shows that the accuracy was close to 90% in both cases. These results indicate that the proposed method can appropriately output findings on road conditions.

Fig 1. Mis-estimated image

4. Conclusion

In this research, we proposed a method for outputting inspection results and findings from images of deteriorated road surfaces using ChatGPT, and for supporting the preparation of reports. We confirmed the usefulness of the proposed method through demonstration experiments. However, when multiple deteriorated conditions are mixed, the proposed method tends to incorrectly estimate the condition of one of the roads. In the future, we aim to improve the estimation accuracy of images with mixed deterioration conditions.

5. Acknowledgment

This work was supported by Hanshin Expressway Young Researcher Grant Fund/Supported Research.

References

Asakawa, H., Nagayama, T., Fujino, Y., Nishikawa, T., Akimoto, T., & Izumi, M. (2012). Development of a simple pavement diagnostic system using dynamic responses of an ordinary vehicle, *Journal of Japan Society of Civil Engineers, Ser. E1 (Pavement Engineering)*, *68*(1), 20–31. https://doi.org/10.2208/jscejpe.68.20.

Imai, R., Nakamura, K., Tsukada, Y., Ito, D., & Kurihara, N. (2021). Research on evaluate cracks in road pavement using drive recorder footage by deep learning, *Journal of Japan Society of Civil Engineers, Ser. F3 (Civil Engineering Informatics)*, *77*(2), I_67–I_76. https://doi.org/10.2208/jscejcei.77.2_I_67.

Digital Life

Journal of Digital Life 2024, B-3

Conference Proceedings

Fundamental Research on Method for Generating Real Scale 3D Shape Reconstruction Results Using Exif Images

Shota Yamashita [1], Masaya Nakahara [1], Yoshinori Tsukada [2] and Yoshimasa Umehara [3]

[1] Faculty of Information Science and Arts, Osaka Electro-Communication University, 1130-70 Kiyotaki, Shijonawate-shi, Osaka 575-0063, Japan
[2] Faculty of Engineering, Reitaku University, 2-2-1 Hikarigaoka, Kashiwa-shi, Chiba 277-8686, Japan
[3] Faculty of Business Administration, Setsunan University, 17-8 Ikedanakamachi, Neyagawa-shi, Osaka 572-8508, Japan

1. Introduction

In recent years, digital twinning has been promoted in Japan to measure real urban spaces and digitize them as 3D models. The main methods used to measure real urban spaces are laser scanner-based measurement and 3D reconstruction from images. Of these measurement methods, the technology for 3D restoration from images cannot determine the actual size and distance in real size on the image. Therefore, manual adjustment to the actual scale is required. However, laser scanners are specialized equipment, and it takes time and experience to be able to measure accurately. Therefore, there is a need for a technology that can easily obtain the actual dimensions of an object using images taken with a smartphone or other device. An existing method for obtaining the actual dimensions of an object in an image is one that can measure the size of the object based on the location information attached to the image and manually selected correspondence points (Kunii, 2013). This method can acquire the actual size of an object using multiple images taken with a smartphone, as in this study. However, since it is necessary to manually specify many corresponding points, it is not possible to obtain the actual dimensions of roads and other objects with few features. Therefore, this study proposes a method to generate a real 3D shape reconstruction result without specifying feature points, based on an image with Exif information including location information taken with a smartphone.

2. Methods

Based on the discussion in Chapter 1, this chapter describes the flow of a method for restoring the generated point cloud data to a 3D shape in real dimensions, using an image taken with a smartphone and its positional information as input. The input data is an image of the measurement target. The location information is set to be saved in the Exif information of the image. Considering the influence of trees and buildings, at the beginning of the image capture, a circular convergence image is taken at a location where it is easy to obtain location information from a GNSS (Global Navigation Satellite System) (Figure 1(a)). The output data is point cloud data in real dimensions. The point cloud data is generated by Gaussian Splatting, a method for recovering 3D shapes from multiple images (Kerbl et al., 2023)). The proposed method first estimates the point cloud data of the subject and its shooting position from the captured image using Gaussian Splatting. Next, the circular trajectory of the estimated shooting position and GNSS position information is obtained, and the radius of the circle is calculated using the least-squares method (Figure 1(b)). The GNSS position information is converted from latitude and longitude to plane rectangular coordinates. The magnification factor of the circle radius of the estimated shooting position relative to the circle radius of the GNSS is calculated and multiplied to the coordinate values of the generated point cloud data (Figure 1(c)). This produces the result of restoring a 3D shape in real dimensions.

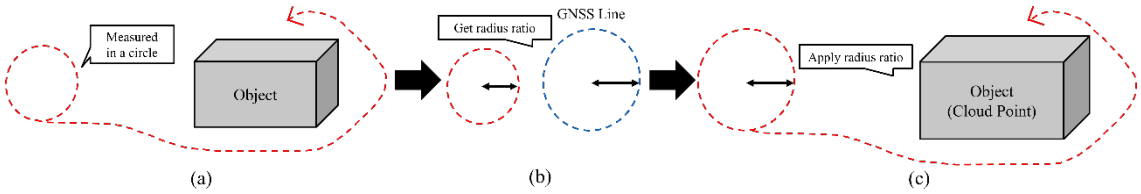


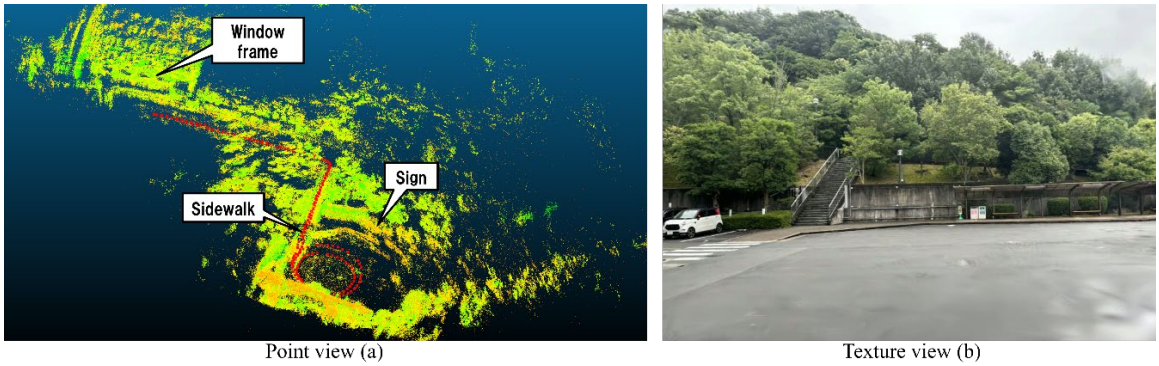
Figure 1 Diagram of the proposed method

Published: 16 September 2024
* Correspondence: nakahara@oecu.jp

3. Demonstration experiment

In this experiment, we evaluate the size and width of geographic features in the point cloud data obtained by the proposed method, using the point cloud data measured by a ground-mounted laser scanner as the most probable value. This confirms the usefulness of the proposed method. In order to verify whether the 3D shape can be reconstructed even if the GNSS satellite signals cannot be received during the measurement due to interference from trees, we selected a location outside a university in the mountains as the measurement site. The ground truth was measured using FARO's “Focus 3D Premium 150. An Apple “iPhone 15” smartphone was used to take 231 photographs to reconstruct the 3D shape. The objects to be evaluated were sidewalk, window frame, and sign, and their widths and heights obtained from the point cloud data of the proposed method were compared with the ground truth.

Figure 2 shows the point cloud data and texture generated by the proposed method, and Figure 3 shows the evaluation results. The red dots in Figure 2(a) are the shooting positions estimated by Gaussian Splatting.



Point view (a) Texture view (b)

Figure 2 Results of point cloud data generation by Gaussian Splatting

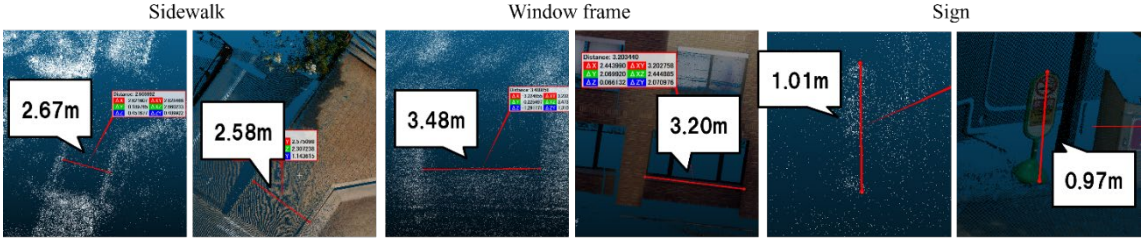


Figure 3 Evaluation results for each object

Figure 3 shows that the proposed method can generate point cloud data that is roughly close to the actual size. In particular, the error was within 0.1m for sidewalk. This result suggests that the proposed method can contribute to solving the problem of not being able to obtain actual size for roads with few features. However, an error of 0.28 m was observed for window frame. This is considered to be because the distance from the shooting position to the evaluation target is far, resulting in a larger error. Because the error varies depending on the geographic feature, further improvement in accuracy cannot be expected simply by changing the magnification factor of the coordinate values of the point cloud data. In addition, in this study, only the dimension values were corrected, and the positions of the point cloud data do not match each other. Therefore, it is necessary to investigate a method to correct the positional relationship and errors of point cloud data generated based on positional information in the future.

4. Conclusions

In this study, we proposed a method for restoring point cloud data generated from images with Exif information including location information to a real 3D shape. We confirmed the usefulness of the proposed method through demonstration experiment. In the future, we will develop a method for correcting the positional relationship and errors in the point cloud data, and aim to improve the accuracy of 3D shape restoration using the proposed method.

References

Kerbl, B., Kopanas, G., Leimkühler, T., & Drettakis, G. (2023). 3D Gaussian splatting for real-time radiance field rendering. *ACM Transactions on Graphics*, *42*(4).

Kunii, Y. (2013). Development of convenient field survey software for historical structure by using a smartphone. *LRJ, 76*(5), 489–492.

Digital Life

Journal of Digital Life 2024, B-4

Conference Proceedings

Fundamental Study on Automated Counting Method of Turning Movement Counts at Intersections Using Video Images

Ryo Sumiyoshi [1], Ryuichi Imai [2], Yuhei Yamamoto [3],
Masaya Nakahara [4], Daisuke Kamiya [5] and Wenyuan Jiang [6]

[1]Doctoral Course Graduate School of Engineering and Design, Hosei University, 2-33 Ichigayatamachi, Shinjuku-ku, Tokyo 162-0843, Japan
[2]Faculty of Engineering and Design, Hosei University, 2-33 Ichigayatamachi, Shinjuku-ku, Tokyo 162-0843, Japan
[3]Faculty of Environmental and Urban Engineering, Kansai University, 3-3-35 Yamate-cho, Suita-shi, Osaka 564-8680, Japan
[4]Faculty of Information Science and Arts, Osaka Electro-Communication University, 1130-70 Kiyotaki, Shijonawate-shi, Osaka 575-0063, Japan
[5]Faculty of Engineering, University of the Ryukyus, 1 Senbaru, Nishihara-cho, Nakagami-gun, Okinawa 903-0213, Japan
[6]Faculty of Engineering, Osaka Sangyo University, 3-1-1 Nakagaito, Daito-shi, Osaka 574-8530, Japan

1. Introduction

In Japan, turning movement counts surveys are conducted to assess traffic conditions at intersections by counting the number of vehicles traveling in each direction (turn right, turn left, and go straight). However, these surveys primarily rely on manual visual observation by surveyors. Due to the necessity of deploying numerous surveyors at each intersection, the associated costs are substantial. Recently, AI-based technologies have gained significant attention for their potential to automate and streamline the data collection process in traffic surveys, thereby reducing the need for manual labor. In AI-assisted surveys, two primary methodologies are employed: the first involves establishing cross-sectional lines around intersections, while the second focuses on analyzing vehicle travel trajectories. However, the method of using cross-sectional lines has a limitation in that counting accuracy may degrade when occlusion occurs as vehicles cross these lines. On the other hand, the method focusing on vehicle trajectories may face challenges in application due to the variability of these trajectories across different road structures. Therefore, the objective of this study is to develop a turning movement counts method that is robust to occlusion, utilizing video images of entire intersections captured by a single fixed-view camera.

2. Methods

The processing flow is illustrated in Fig. 1. This method utilizes video images that captures the entire intersection as input and outputs the turning movement counts. First, in the process of setting cross-sectional lines, four lines (L1, L2, L3, L4) are established to encircle the intersection (Fig. 2). This allows for the determination of the direction from which vehicles enter and exit the intersection. In the auxiliary line setting process, auxiliary lines are established by connecting the midpoints of opposite edges to mitigate the effects of occlusion. This approach enables the determination of the direction from which a vehicle enters the intersection, even in cases where occlusion occurs during entry. Next, in the vehicle detection process, the YOLOv8x model, an advanced deep learning-based object detection method, is applied for vehicle detection. In the vehicle tracking process, the BoT-SORT algorithm is employed to track the detected vehicles. Finally, in the counting process, vehicles are counted as one unit when they pass through two section lines or auxiliary lines. The process then outputs the number of vehicles that traveled in each direction.

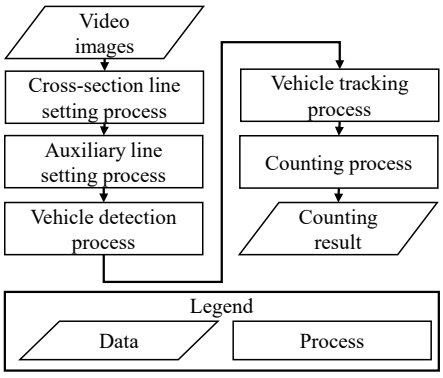


Fig. 1. Processing Flow

Published: 16 September 2024
* Correspondence: imai@hosei.ac.jp

3. Experiment and Results

In this chapter, we apply the proposed method to video images captured by a camera positioned to cover the entire intersection and evaluate the method's effectiveness. The video images used for verification was recorded at a resolution of 4K and a frame rate of 30 fps. For the evaluation, precision, recall, and F-measure are calculated for all directions, and the average of each metric is used. The experimental results are presented in Table 1. Table 1 indicates that the average F-measure exceeded 0.978 in all directions. Since the accuracy of manual observation is estimated to be around 95.0%, the counting accuracy obtained in this experiment can be considered equivalent to or better than that achieved through manual counting. These results confirm the effectiveness of the proposed method.

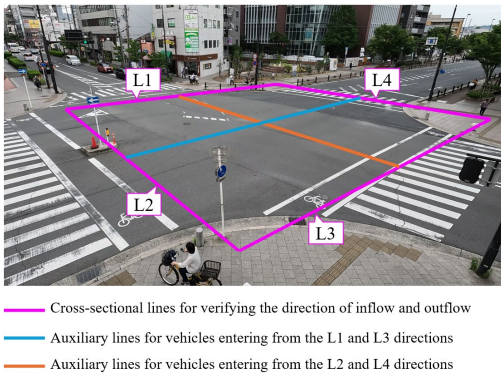


Fig. 2. Example of Cross-Sectional Line Setting

Table 1. Directional Traffic Count Results

Inflow direction	Ground Truth (vehicles)	Predicted Count (vehicles)	True Positive Count (vehicles)	Average precision	Average recall	Average F-measure
L1	137	139	137	0.990	1.000	0.995
L2	54	51	51	1.000	0.957	0.978
L3	61	61	61	1.000	1.000	1.000
L4	126	127	126	0.989	1.000	0.994

4. Discussions

Detailed examination of the experimental results revealed that vehicles entering from the L2 direction and exiting in the L4 direction were correctly counted, despite instances where the ID assigned during vehicle tracking switched due to occlusion. This was made possible by setting a cross-sectional line connecting the midpoints of the opposing edges, which allowed for accurate counting even when occlusions occurred within the rectangular area. However, as shown in Fig. 3, vehicles entering from the L2 direction and exiting in the L1 direction experienced ID switches due to occlusion, which led to counting omissions. As a countermeasure, attention could be directed to the vehicle's travel path within each of the four rectangles formed by the set cross-sectional lines. Specifically, turning left vehicles, regardless of the direction they enter from, have a travel trajectory that remains within a single rectangle. Considering this characteristic in the counting process could lead to improved accuracy.

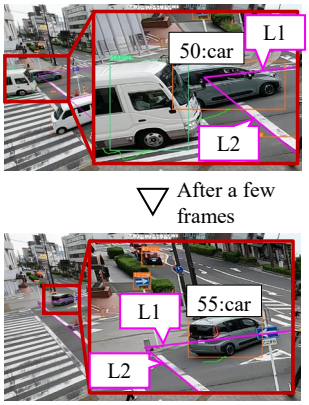


Fig. 2. Vehicles Omitted from the Count Due to Occlusion

5. Conclusions

In this study, we devised a turning movement counts method that is robust to occlusion, using video images captured by a single camera covering the entire intersection. The experimental results indicate that the devised method achieved an accuracy comparable to or greater than manual observation, as the F-measure exceeded 0.978 in all directions. Future work will focus on improving the counting accuracy for turning left vehicles. In addition, we will develop a method for classifying the types of counted vehicles. Additionally, since the video images used for this verification was from a road with two lanes in each direction, we will examine whether the same level of accuracy can be achieved when the method is applied to a road with three lanes in each direction.

References

Aharon, N., et al. (2022). BoT-SORT: Robust associations multi-pedestrian tracking. https://arxiv.org/abs/2206.14651

Horii, D., et al. (2022). A study on automatic measurement of precise traffic engineering indicators volume by intersection direction using deep learning, *Intelligence, Informatics and Infrastructure*, *3*(J2), 819–825.

Jocher, G., et al. (2023). Ultralytics YOLOv8, https://github.com/ultralytics/ultralytics

Watanabe, K., et al. (2023). Development and its evaluation of a counterline optimization method for directional traffic surveys using MOT, *Transactions of Information Processing Society of Japan, 64*(2), 511–520.

Digital Life

Journal of Digital Life 2024, B-5

Conference Proceedings

Prototype of Digital Twin Environment Using 3D Spatial Information Infrastructure Based on Spatial IDs

Kenji Nakamura [1], Yoshimasa Umehara [2], Masaya Nakahara [3], and Ryuichi Imai [4]

[1] Faculty of Information Technology and Social Sciences, Osaka University of Economics, 2-2-8 Osumi, Higashiyodogawa-ku, Osaka 533-8533, Japan
[2] Faculty of Business Administration, Setsunan University, 17-8 Ikedanakamachi, Neyagawa-shi, Osaka 572-8508, Japan
[3] Faculty of Information Science and Arts, Osaka Electro-Communication University, 1130-70 Kiyotaki, Shijonawate-shi, Osaka 575-0063, Japan
[4] Faculty of Engineering and Design, Hosei University, 2-33 Ichigayatamachi, Shinjuku-ku, Tokyo 162-0843, Japan

1. Introduction

With the promotion of digital transformation (DX) in various industries, the demand for digital twins (Madni et al., 2019), including geospatial information, is increasing. Recently, a social trend toward the openness of related information has been observed on platforms such as the National Land Transportation Data Platform and PLATEAU. In addition, the G-Spatial Information Center enables cross-disciplinary searches for these data, creating an environment in which the data necessary for constructing new businesses and services related to spatial information can be easily obtained. However, challenges remain, such as the proliferation of platforms requiring service providers to search for necessary data, lack of an environment for instantly obtaining spatial information around specific locations, need to analyze raw data as necessary, and necessity of developing systems for data access and analysis that accommodate various data formats. To address these problems, constructing a three-dimensional (3D) spatial information infrastructure with a common index, the spatial ID (METI et al., 2022), is necessary, which allows cross-disciplinary access to data held by various business entities and facilitates machine-to-machine communication. Therefore, this study aims to develop a socially implementable 3D spatial information infrastructure based on the technical specifications of spatial IDs and a prototype digital twin environment.

2. Overview of the 3D Spatial Information Infrastructure

Various spatial information such as human flow, weather data, and inspection information of public structures are generally managed using latitude, longitude, and altitude. However, the traditional method faces problems, such as the complexity of managing detailed coordinate values and difficulty in identifying parts of complex structures such as bridges. To resolve these problems, we developed a 3D spatial information infrastructure that manages spatial information based on spatial IDs and assigns unique IDs to voxels, thus dividing the 3D space into fixed sizes. As shown in Fig. 1, this infrastructure manages dynamic data, such as weather and human flow, as well as static data, such as map and point cloud data, using a blockchain based on spatial IDs. By constructing this 3D spatial information infrastructure with spatial IDs as a common index, we achieved a digital twin that can be used cross-disciplinarily in various industries for location information.

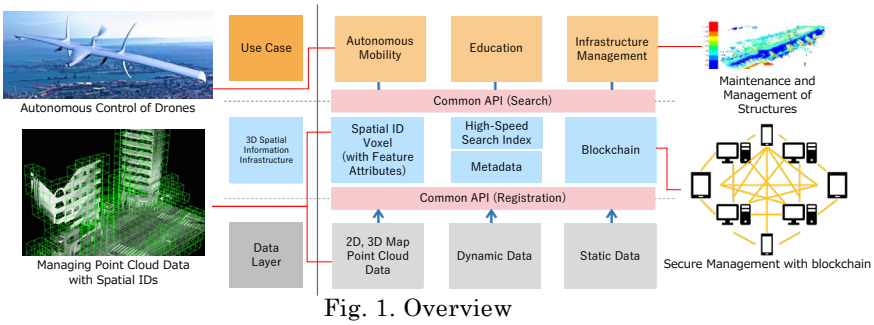


Fig. 1. Overview

Published: 16 September 2024
* Correspondence: k-nakamu@osaka-ue.ac.jp

3. System Architecture

The 3D spatial information infrastructure shown in Fig. 2 comprises a construction server and a management server. The construction server automatically generates spatial voxels for each road structure from the point cloud, map, and PLATEAU data. The management server maintains dynamic data, such as weather and human flow data, associated with these spatial voxels in the 3D spatial information infrastructure. Users can access the 3D spatial information infrastructure through the external application programming interface of the management server. By utilizing the digital twin constructed with this 3D spatial information infrastructure, it is easy to perform interference checks with structures on a route, set optimal routes based on dynamic data for autonomous drone navigation, efficiently maintain public structures during normal times, and immediately identify damaged areas during disasters in infrastructure monitoring. Additionally, data scientists can be trained to perform spatial and human flow data analyses using digital twins in future DX societies.

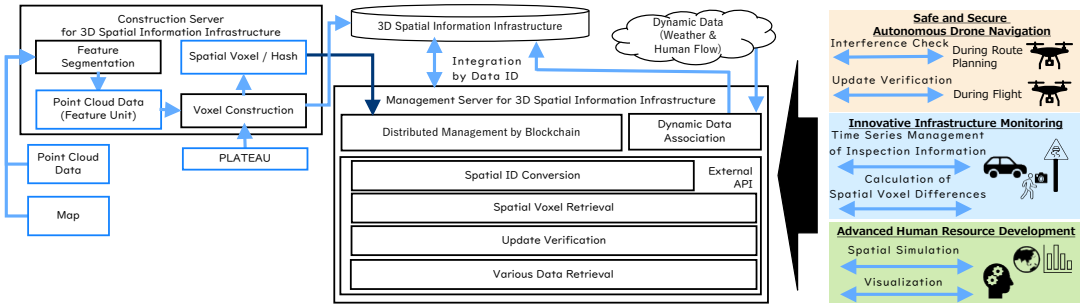


Fig. 2. System architecture

The features of the 3D spatial information infrastructure are as follows:

- The high-speed search index reduces the time required for reference and comparison by 1/100th of the traditional time.
- Spatial IDs facilitate the easy association of data with 3D structures and specific parts of these structures.
- Blockchain technology ensures traceability and secure management of vast and diverse spatial information.
- Point cloud data are automatically divided into six types of structures (roads, railways, waterways, buildings, towers, and power lines) and allow the real-time retrieval of any data according to the application needs (Fig. 3).
- Spatial voxels are constructed in 1,200 map sections of Shizuoka Prefecture, achieving bit operations in approximately 3s or less in any space (Fig. 4).

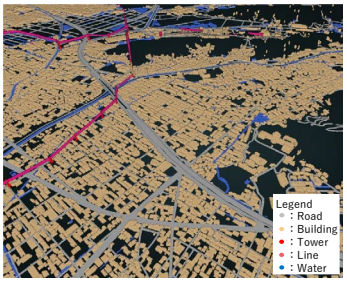


Fig. 3. Digital twin environment in Shizuoka Prefecture

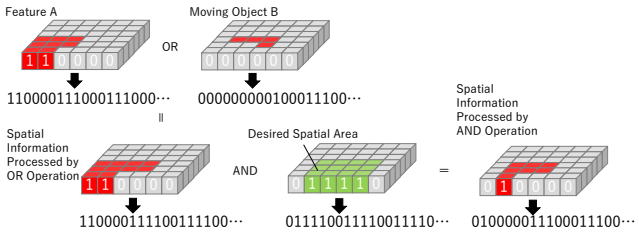


Fig. 4. Bit operation image of voxel space

4. Conclusion

In this study, we developed a 3D spatial information infrastructure based on the technical specifications of spatial IDs, and prototyped a digital twin environment in Shizuoka Prefecture, Japan. Using spatial IDs, we enabled cross-disciplinary searches and analyses and successfully constructed a high-speed data distribution infrastructure. In the future, we plan to apply and evaluate this infrastructure for various use cases and make improvements based on feedback during practical use.

Acknowledgments:

This study is based on the results obtained from a project commissioned by the New Energy and Industrial Technology Development Organization (NEDO).

References

Madni, A., Madni, C., and Lucero, S. (2019). Leveraging digital twin technology in model-based systems engineering, *Systems*, *7*(7). https://doi.org/10.3390/systems7010007

METI and DADC (2022). *3D Spatial Information Infrastructure Architecture Design Report*.

Digital Life

Journal of Digital Life 2024, B-6

Conference Proceedings

Application of machine learning to single-cell RNA sequencing provides the effective drugs against drug-tolerant persister cells in colorectal cancer

Yosui Nojima [1,*], Ryoji Yao [2] and Takashi Suzuki [1]

[1] Center for Mathematical Modeling and Data Science, Osaka University
[2] Department of Cell Biology, Cancer Institute, Japanese Foundation for Cancer Research

1. Introduction

The inactivation of the *APC* gene is a critical early event in the development of colorectal cancer (CRC). Familial adenomatous polyposis (FAP) is a hereditary syndrome characterized by the development of numerous adenomas in the colon, leading to a high risk of CRC. It is caused by an autosomal dominant mutation in the *APC* gene. The carcinogenesis mechanism of FAP aligns with that of sporadic CRC.

We previously evaluated the efficacy of anticancer drugs using an organoid culture system derived from benign and malignant tumors of FAP patients (Sakahara et al., 2019). The results revealed that organoids derived from malignant tumors are resistant to the MEK inhibitor trametinib. It is known that cancer tissues contain drug-tolerant persister (DTP) cells, which are resistant to drugs (Dhanyamraju et al., 2022).

Single-cell RNA sequencing (scRNA-Seq) is an emerging technology that measures the gene expression levels for each transcript within each individual cell of the sample and allows a representation of the distribution of this expression in each subpopulation of cells and provide a new opportunity to probe cancer biology in a high-resolution and dynamic manner.

Machine learning (ML) is a branch of artificial intelligence that exploits many statistical techniques to allow mathematical functions to learn from experience acquired from similar example. ML methods are increasingly used in cancer diagnosis, prognosis and treatment guidance; especially with scRNA-Seq data as inputs to feed linear models or nonlinear models.

In this study, we performed scRNA-Seq on organoids of FAP and built ML models using public data to identify DTP cells within the organoids being resistant to MEK inhibitors. This study is the first to demonstrate that DTP cells can be identified using ML models. Furthermore, we identified drugs effective against DTP cells in the organoids of FAP using public drug sensitivity data. This study proposes a novel strategy for identifying DTP cells and suggesting effective drugs against the cells.

2. Methods

FAP tumor samples were obtained from consenting patients, and all procedures were approved by the Research Ethics Board at the JFCR Cancer Institute. The tumors were cut into small pieces, enzymatically dissociated, and the resulting cell pellets were suspended in Matrigel. Organoids were cultured in ENR medium at 37°C and 5% O2, with media changes every three days.

For scRNA-Seq, organoids were dissociated, and single-cell libraries were prepared using the Chromium Single Cell 3’ Solution Reagent Kits. Sequencing was performed using a HiSeq2500 system, and the data were processed with Cell Ranger, velocyto, and Seurat packages. Cells were filtered based on mitochondrial gene expression and unique feature counts, and the data were integrated and analyzed to visualize cell clusters.

Published: 16 September 2024
* Correspondence: nojima.yosui.mmds@osaka-u.ac.jp

Machine learning models were constructed to classify DTP cells using integrated a public scRNA-Seq data and FAP data. Various algorithms, including support vector machines and random forest, were applied, and optimal hyperparameters were identified through grid search. The classification models were used to predict DTP cells in the FAP scRNA-Seq data. Drug sensitivity data were sourced from PRISM, CTD[2], and GDSC.

3. Results

We conducted scRNA-Seq on three samples: organoids from benign tumors, malignant tumors, and normal tissues from sporadic CRC patients. Based on marker gene expression, we annotated nine cell types. Among these, two cell clusters unique to tumor-derived organoids were identified and defined as TC1 and TC2. Notably, TC1, prevalent in malignant tumor organoids, expressed cancer stem cell marker genes. Next, we obtained scRNA-Seq data for DTP and non-DTP cells from public databases and constructed a classification model. The model identified approximately 35% of TC1 as DTP cells, while other cell types were below 10% (**Fig. 1 A and B**). We then applied the model to classify cell lines from public gene expression data into DTP and non-DTP cells. Effective drugs for DTP cells were obtained from public databases and tested on malignant tumor organoids, confirming a synergistic effect with trametinib.

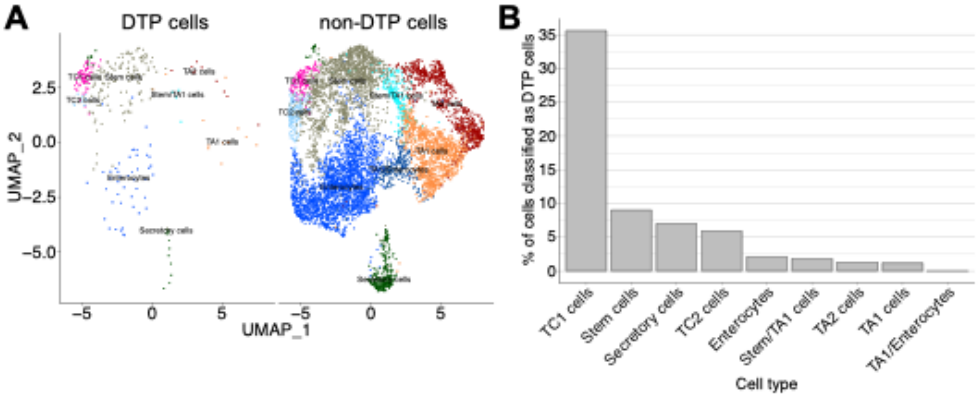


Fig. 1. Classification of single-cell in FAP organoids using a svmRadial model. **A** Cells classified as either DTP or non-DTP cells are projected into a UMAP-based embedding. **B** Percentage of cells classified as DTP cells in each cell type of FAP.

4. Discussions

In this study, scRNA-Seq was performed on organoids derived from benign, malignant, and normal tissues. TC1 and TC2 cell clusters were prominent in tumor organoids. Upregulated genes in TC1 included *PERP*, *TESC*, *TGFBI*, *TMEM59*, and *SOX4*, which are associated with cancer stem cells. To investigate whether TC1 cells are DTP cells, ML models were built using a public scRNA-Seq dataset. The ML model classified over 35% of TC1 cells as DTP cells. Thus, drug resistance in malignant organoids was attributed to TC1 cells, prompting the search for effective drugs. Using DepMap, CCLE, and COSMIC databases, 22 cell lines were analyzed, and three drugs were validated as candidate drugs effective against TC1 cells. Of the three drugs, two drugs showed a synergistic effect with trametinib. This study demonstrates that DTP cells are not resistant to all drugs and that it is possible to propose effective drugs.

5. Conclusions

This study constructed a high-performance ML model for DTP cell identification in FAP scRNA-Seq data and highlighted the synergistic effect of trametinib and the candidate drugs against colorectal cancer cells in FAP patients.

References

Dhanyamraju, P. K., Schell, T. D., Amin, S., & Robertson, G. P. (2022). Drug-tolerant persister cells in cancer therapy resistance. *Cancer Research*, *82*(12), 2503–2514. https://doi.org/10.1158/0008-5472.can-21-3844

Sakahara, M., Okamoto, T., Oyanagi, J., et al. (2019). IFN/STAT signaling controls tumorigenesis and the drug response in colorectal cancer. *Cancer Science*, *110*(4), 1293–1305. https://doi.org/10.1111%2Fcas.13964

Digital Life

Journal of Digital Life 2024, B-7

Conference Proceedings

Dynamic Control of the Optimal Product Prices Using Machine Learning

Kazuhiro Miyatsu

School of Information and Network, Senshu University

1. Introduction

Dynamic pricing has been studied as revenue management in hotel and airline industries from operation research viewpoints by predicting demand in the market. In the digitalized society with big data, dynamic pricing is expanding to include consumer goods analyzing with advanced statistical models (Miyatsu, 2023). The purpose of the research is to design a personalization framework of the optimal prices for commodity goods using machine learning.

2. Methods

Machine learning technology deployed in the models is XGBoost in the two steps: the first model is to estimate product purchase probability for actual given prices from each store (*ML_1*), and the second is to predict prices with probability estimated ($\overline{buy}$) from the first model and actual product prices as labels (*ML_2*). The second model is then used to dynamically control the optimal product prices ($price^*$) for a particular customer with a specific purchase probability (buy^*) in the Fig.1 as well as modeling features and labels in Table.1 In general, classification problem in machine learning is to determine if an incident occurs or not depending on probability value generated as outputs, but in this machine learning model (*ML_2*) estimated prices are conditional to purchase probability. $buy^*(\%)$ dynamically controls the optimal product price $price^*$(JPY) for a particular customer, where $buy^*(\%)$ is critical value to calculate expected revenue from the optimal prices offered $price^*$(JPY).

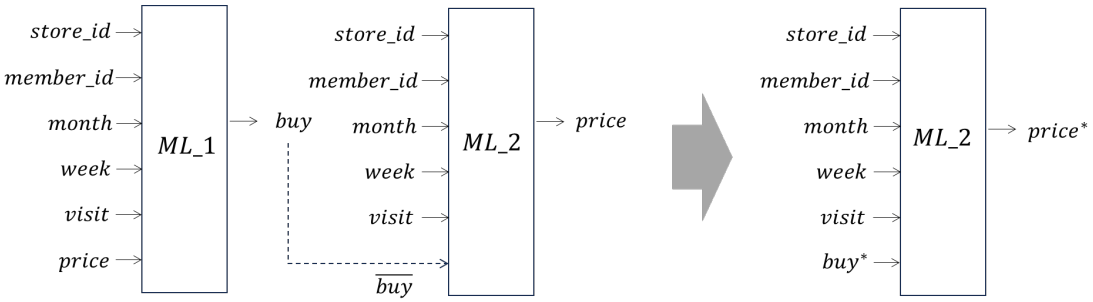


Fig. 1. The Optimal Product Price Prediction Models

Table 1. Features and labels deployed in tuning XGBoost models

Features & Labels	Description	Value
store_id	The store number where the customer usually makes purchases	79 bits binary
member_id	The unique customer identification number	569 bits binary
month	The monthly identification number	12 bits binary
week	The weekly identification number	7 bits binary
visit	The store visit flag of the customer	2 bits binary
buy	The product purchase flag of the customer	2 bits binary
price	The product price of a store where the customer belongs	numeric

Expected revenue ($E[Revenue_i^*]$) from the optimal product prices is formulated in Equation (1). N is the number of estimations during a purchase period and $visit_{i,t}$ is an indicator where it is set to 1 when a customer(i) visits the store at the time(t). We do not necessarily have to know exactly when product purchase occurs but shall know

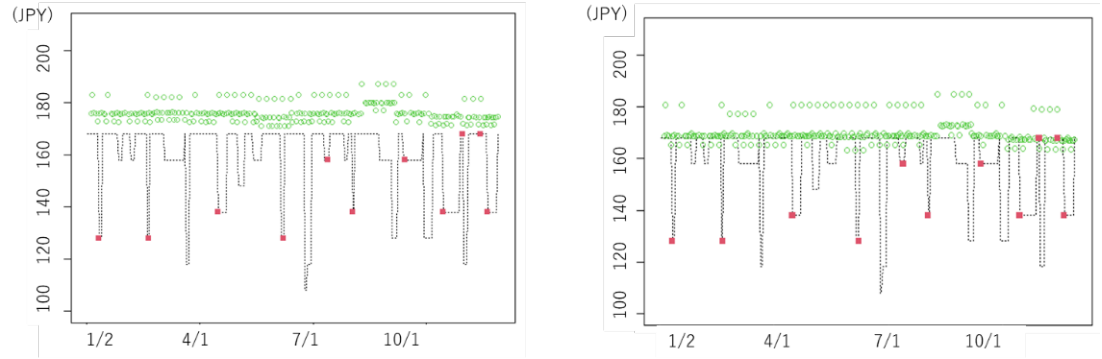
Published: 16 September 2024
* Correspondence: miyatsu@isc.senshu-u.ac.jp

likelihood of purchase for the prices, and its expectation value is to be compared with actual revenue during an observation period for evaluation of the proposed pricing strategy.

$$E[Revenue_i^*] = buy_i^* \times \left\{ \sum_{t=1}^{N} price_{i,t}^* \times visit_{i,t} \right\} \quad i = 1,2,3, \ldots \ ,596 \tag{1}$$

3. Results

To develop the XGBoost models, i.e. *ML_1* and *ML_2* depicted in Fig.1, customer purchase history data during a period of 2019/1/2-2019/12/31 at retail stores in Tokyo metropolitan area is used. A product selected for analysis is ketchup (KAGOME500G), and its prices differ by stores even on the same date. To specify product prices of each store, 79 stores identification is deployed as well as 596 members identification. To estimate the optimal prices for customer(i), product purchase likelihood must be set as $buy_i^*(\%)$. Fig.2 (a) and (b) show the optimal product prices, as an example, with two different probabilities, i.e. buy_i^* = 4.25% and 4.63%. The dotted lines are selling prices, whose normal price is 170 (JPY), but it occasionally displays dips as promotional discount. The red dots are actual purchase points of this particular customer, and green ones are the optimal prices suggested from the model. In Fig.1 (a), it suggests higher prices than those in (b) as purchase probability is set lower than that in (b). Price elasticity is normally negative to its own price, so that prices become higher when likelihood is set lower, and vice versa. For this case, the customer bought products 11 times out of 259 store visits during a period, which accounts to 4.25% purchase likelihood. The model suggests slightly higher prices for such shopping pattern as shown in Fig.1 (a), and it also suggests lower prices when buying frequency is assumed to increase up to 6.50%. For 2 (a) and(b) cases, based on equation (1), expected revenue increase 22.0% and 28.4% respectively. Such analysis is valid for the rest of other customers. However, the purchase patterns and prices are heterogenous among customers, so efficiency of the optimal prices derived from the model differs. Overall expected revenue increase is approximately 10% by properly selecting customers to apply for the results.



(a) buy_i^* = 4.25 (%) (a) buy_i^* = 4.63 (%)

Fig.2. The Optimal Product Prices of one Customer

4. Discussions

Altering likelihood of purchase shall be cautious as it would impose shopping pattern changes. Increasing buy_i^* from 4.25 to 4.63 implies an additional product purchase for the case described in the previous section. Considering practicality, it would be reasonable to increase buy_i^* value up to 10% to avoid changing each customer's shopping trait. As for features deployed in the models, the current features are seasonality-oriented binary except for product prices. In addition to such binary features, customer-oriented variables such as amount spent at the previous shopping occasion and/or the number of days elapsed since the last purchase would be worth being included to capture individual shopping characteristics.

5. Conclusions

The optimal product pricing control method using machine learning is proposed as dynamic personal pricing strategy in marketing, and the models were developed using XGBoost algorithm with individual purchase history records of 596 customers from 79 stores during a period of 2019/1/2-2019/12/31.Overall revenue increase can be expected around 10% by setting proper purchase rate to the model.

References

Miyatsu, K. (2023). Dynamic pricing strategies: The optimal prices for consumers and channels, *Network and Information Review of Senshu University*, 11–18.

Miyatsu, K. (2023). Dynamic price personalization: Predicting the optimum price to maximize revenue, *IEEE International Conference on Big Data Proceeding*, 3439–3444.

Digital Life

Journal of Digital Life 2024, B-8

Conference Proceedings

Empowering Educators: Assessing and Advancing Minecraft Education Edition Resources

Yan Zhu [1], Ryousuke Furukado [1]

[1] Nishinippon Institute of Technology, 1-11, Aratsu, Miyako-gun, Fukuoka, 800-0394, Japan.

1. Introduction

In recent years, the use of digital tools in education has been rapidly advancing. In particular, digital games have potential as educational tools, among which Minecraft: Education Edition (MEE, Microsoft, 2016) is gaining attention as an attractive learning tool for both educators and students. MEE is a real-world element and provide many benefits to teachers, parents, and administrators by enhancing creativity and promoting community interaction (Sánchez-López et al., 2022). MEE is also a platform for educators to provide interactive and creative learning experiences for students, especially in STEM education.

MEE is primarily intended for use in educational settings, where educators can use MEE to teach a wide variety of subjects, including coding, science, history, and art. However, in order to use MEE effectively, it is essential that educators have the proper resources and training. Therefore, Microsoft has created a Teacher Academy (TA) program for educators to provide training programs and educational resources related to MEE. However, despite the growing popularity of MEE as an educational tool and its great potential, Microsoft's TA program offerings remain limited. Existing research highlights the positive impact of MEE on student engagement, creativity, collaboration, and problem-solving skills, but there is little research specific to teacher pedagogy.

Therefore, the purpose of this study was to compare and evaluate the resources provided to educators by the TA program for MEE with existing educational resources and to propose specific improvements based on educators' needs.

2. Methods

The following table summarizes the TA as of June 1, 2024 (Table 1).

Table 1: Overview of TA

Modules	Overview	Required Time
Minecraft 101: Let's build individually	Minecraft Education supports game-based teaching and learning in a creative, collaborative, and problem-solving immersive environment. Model a house building lesson, assess the lesson, and apply it in the classroom with a lesson planner.	1 hour 10 minModule 8 Units
Minecraft 201: Let's modify together	Build upon the skills learned in Minecraft 101 by continuing another build challenge introducing new in-game assessment tools and a multiplayer environment. Continue to look into the lesson planning process as a reflection.	1 hour 3 minModule 6 Units
Minecraft 301: Let's build worlds together	Build on the skills in Minecraft 101 and 201 through a build challenge with a small-group multiplayer environment that introduces world builders and special blocks. Continue to look at lesson planning as reflection.	1 hour 10 minModule 6 Units

Next, existing e-learning education programs are surveyed, and specific examples are presented in the results. In addition, differences between existing educational programs and TA programs will be extracted and problems will be outlined.

Published: 16 September 2024
* Correspondence: s240102@nishitech.ac.jp

3. Results & Discussions

MEXT Japan's School Safety e-Learning for Teachers and Staff (2024) program is an online program for teachers to learn about safety management and crisis response in the school setting, enabling them to acquire practical skills through specific examples and simulations. The program strikes a balance between theoretical and practical elements, thereby enabling teachers to enhance their knowledge and responsiveness to real-world challenges. On the other hand, the MEE TA program specializes in teaching methods using MEE, but it is limited in the acquisition of specific practical skills, making it difficult to fully develop the ability to respond to the diverse issues faced in actual educational settings. Another problem is the lack of training in realistic scenarios through simulations such as those provided by the School Safety e-Learning program, which leaves teachers ill-prepared to respond quickly and appropriately in realistic situations.

The University of Tokyo's Faculty Development (Tokyo FD) (2024) program aims to provide faculty members with opportunities to learn diverse educational methods and theories to improve their teaching skills. The Tokyo FD offers training on a diverse array of pedagogical approaches, including active learning and flipped classrooms. This enables faculty members to develop the skills required to translate educational theory into practice. The evaluation methods are also diverse and reliable, and teachers are supported in effectively assessing student learning outcomes. In contrast, the MEE TA program is specialized in MEE-based teaching methods, which limits the diversity of teaching methods and provides few opportunities to incorporate other teaching techniques and methodologies. Furthermore, the program lacks sufficient support and feedback systems for long-term educational improvement and lacks support for faculty members to continuously improve their own teaching methods. Therefore, MEE's TA program is inferior to the University of Tokyo FD program in terms of understanding educational theory and continuous educational improvement.

The MEE TA program differs in its achievement goals from general e-learning programs for educators because it aims to enable educators to use games on MEE to teach and develop their own lesson plans. Therefore, MEE TA programs should include more MEE-related knowledge and methods to facilitate courses. For example, suggestions could include how to use the command box in the game and how to use commands to quickly construct large buildings. And several shortcomings can be noted in the resources provided by the current MEE TA program. In particular, the main challenges faced by educators include the following(Table 2).

Table 2: Current MEE TA Program Shortfalls

item name	Contents
Lack of quality and quantity of resources	The materials and support materials provided are limited and do not address all learning levels and educational environments.
Lack of technical support	Lack of adequate support for technical problems and questions prevents educators from using MEE with confidence.
Lack of practical training	Lack of specific training for educators to effectively use MEE in their actual classes.
Unclear evaluation criteria	Criteria and rubrics for assessing the outcomes of learning activities using MEE are unclear.
Lack of multicultural support	Insufficient support for students from different cultural backgrounds.

4. Conclusions

Therefore, it will be necessary to aim to update the MEE TA program by having in-service teachers take the current TA program. Additionally, a questionnaire survey should be conducted to clarify its applicability to educational practice, the level of enrichment and satisfaction with technical support, and to gather suggestions for improvement from an educator's perspective.

References

Microsoft. (n.d.). Minecraft Education: Teacher Academy. https://learn.microsoft.com/en-us/training/paths/minecraft-teacher-academy/

Ministry of Education(n.d.). School Safety E-Learning for Teachers and Staff. https://anzenkyouiku.mext.go.jp/learning/index.html

Sánchez-López, I., Roig-Vila, R., & Pérez-Rodríguez, A. (2022). Metaverse and education: the pioneering case of Minecraft in immersive digital learning. doi: 10.3145/epi.2022.nov.10

The University of Tokyo(n.d.). Interactive Teaching. https://www.coursera.org/learn/interactive-teaching

Digital Life

Journal of Digital Life 2024, B-9

Conference Proceedings

What motivates evacuees to leave their cars during a tsunami evacuation?

Toshiya Arakawa [1]

[1] Department of Data Science, Faculty of Advanced Engineering, Nippon Institute of Technology

1. Introduction

Investigations conducted after the Great East Japan earthquake of 2011 reveal that more than 60% of evacuees used automobiles to escape from the disaster zone (Murakami et al., 2012), which resulted in a severe gridlock in central Ishinomaki, Miyagi Prefecture, Tohoku region (Hara and Kuwahara, 2015). Using private cars to evacuate city areas during a disaster can result in secondary disasters. Evacuees who did not leave their car during the 2011 tsunami evacuation suffered significantly. However, research has not yet identified what motivates evacuees to leave their cars. Therefore, we developed a tsunami evacuation experience simulator, conducted an experiment, and determined the factors influencing evacuees to leave their cars during a tsunami evacuation.

2. Methods

We developed a tsunami evacuation experience simulator (Figure 1(a)) comprising a desktop computer, a steering controller, and a smartphone (iPhone 14, Apple Inc.) based on the simulator–platform software Sirius (Misaki Design LLC). Sirius uses Unreal Engine 4 (Epic Games) for real-time computer graphics rendering and the open-source multi-agent traffic simulation program Re:sim (Misaki Design LLC) to reproduce the required traffic environments, such as traffic jams and evacuees on foot (including children, older adults, and wheelchair-using individuals).

In the experiment, the smartphone was used to detect participants' motion, and the acquired motion data were then transferred to the wi-fi. The tsunami evacuation experience simulator simulated both car and on-foot evacuation scenarios. First, we simulated the evacuation by car scenario using the steering controller (Figure 1(a)). The participant driving the car could choose to leave their car and evacuate on foot at any time. To leave their car in the simulation, the participant had to stand up from their seated position; this motion would be detected and captured by the smartphone camera, and the evacuation mode would change to the on-foot scenario. To detect the driver's motion, we used ThreeDPoseTracker (Digital Standard Inc., 2024), which is an iOS app that detects the 3D coordinates of joints in the human body using camera images. Figure 1(b) presents the evacuation on foot scenario. During the evacuation by car scenario, participants view the simulation from the driver's viewpoint. However, once the participants choose to leave the car, they are able to see the avatar of the driver/evacuee moving in the simulation environment.

Fig. 1. Tsunami evacuation experience simulator. (a) evacuation by car, (b) evacuation on foot.

Published: 16 September 2024
* Correspondence: arakawa.toshiya@aut.ac.jp

In the experiment, the simulator can choose to display the tsunami arrival time, the distance to the evacuation site, and whether the driver of the car in front of the focal participant's vehicle leaves their car.

3. Results

Table 1 shows the number of participants who chose to leave their cars. Nine of the 27 participants, or approximately one-third of the participants, left their cars at least once to evacuate on foot. We used Hayashi's Quantification Method II, which is a discriminant analysis program based on qualitative data. The objective variable was whether the driver left the vehicle, and the explanatory variables were whether the tsunami arrival time was displayed, whether the distance to the evacuation site was displayed, and whether the person in the car in front of the focal participant left their car. Table 2 shows the linear discriminant coefficients obtained using Hayashi's Quantification Method II.

Table 1. The number of participants who chose to leave their cars.

	When the person in the car in front of the focal participant leaves the car.	When the person in the car in front of the focal participant does not leave the car.
Both time and distance are displayed	4	2
Only time is displayed	4	0
Only distance is displayed	4	1
None	3	1
Total	15	4

Table 2. The linear discriminant coefficients obtained using Hayashi's Quantification Method II

Explanatory variables	Coefficients of linear discriminants
Whether the tsunami arrival time is displayed	0.177
Whether the distance to the evacuation site is displayed	0.531
Whether the person in the car in front of the focal participant leaves the car	1.948

4. Discussion

Table 1 shows that when people in front of the focal participants' vehicles left their cars, 15 participants left their vehicles and four did not. While the total number of people who left their vehicles was small, the simulation results suggested that during an evacuation, people may be more willing to leave their cars when they see people in the vehicles in front of theirs get out to continue on foot. Table 2 also confirms this supposition as the coefficients for whether a person is willing to leave their car were the highest when the person saw other people leave their cars. Therefore, considering the principle of evacuation on foot, it is suggested that people who decide to leave their cars can set off a chain reaction that motivates other car evacuees to do the same.

5. Conclusion

To verify what motivated evacuees to leave their cars during a tsunami evacuation, we developed a tsunami evacuation experience simulator. Using this simulator, we found that car evacuees are influenced to leave their car when they see other drivers doing so. When a person leaves their car, they set off a chain reaction in which other people watching them are motivated to do the same. Therefore, evacuees can be advised to leave their cars if they get stuck in a traffic jam during an evacuation, which is expected to set off a chain reaction.

References

Digital Standard Inc. (2024). TDPT (iOS ver.), Retrieved from https://digital-standard.com/tdptios_en/ (in Japanese) (accessed on June 1, 2024).

Hara, Y., & Kuwahara, M. (2015). Traffic monitoring immediately after a major natural disaster revealed by probe data - A case in Ishinomaki after the Great East Japan Earthquake, *Transp. Res. Part A Policy Pract. 75*, 1–15. https://doi.org/10.1016/j.tra.2015.03.002

Murakami, H., Takimoto, K., & Pomonis, A. (2012). Tsunami evacuation process and human loss distribution in the 2011 Great East Japan Earthquake - A case study of Natori City, Miyagi Prefecture, In *Proceedings of 15th World Conference on Earthquake Engineering*. 1–10.