Development of Monocular Vision-Based Tracking Method for Wheelchair Sports

Shimpei Aihara[1][a], Takara Sakai[2] and Akira Shionoya[2]
[1]Department of Sport Science and Research, Japan Institute of Sport Sciences, Tokyo, Japan
[2]Department of Management and Information Systems Science, Nagaoka University of Technology, Niigata, Japan

Keywords: Wheelchair Sports, Positioning System, Monocular Camera, Deep Learning.

Abstract: Recently, tracking systems to measure player positions have been introduced in the sports domain. However, wheelchair sports have not been considered extensively. In addition, user-friendly and low-cost systems for wheelchair sports are uncommon. Thus, in this paper, we propose a method to calculate the kinematic data of wheelchair athletes on a playing field (i.e., player positions and wheelchair directions) using images acquired by a monocular camera. The proposed method was evaluated experimentally, and the root mean square error of the position accuracy was 0.11 m, and the mean average error of the direction accuracy was 6.78 degrees. The results demonstrate that the proposed method outperforms existing tracking methods in terms of accuracy. The findings of this study suggest that it is possible to acquire kinematic data of wheelchair athletes using a simple method, which we expect to contribute to improvement analysis of the wheelchair athlete performance.

1 INTRODUCTION

Sports promote mental and physical development, enrich humanity, and play an important role in living a healthy life. For people with disabilities, sports can be an important component of their medical rehabilitation. In addition, sports can provide lifelong recreation and can be played at various levels including highly competitive ones, e.g., the Paralympics. In Japan, particularly in competitive sports, interest in sports for the disabled increased due to success of the Tokyo 2020 Olympic and Paralympic Games. Wheelchair sports accounted for about 50% of the competitions held at the Paralympic Games, e.g., tennis, basketball, athletic sports, badminton, rugby, and table tennis. Wheelchair sports are recognized as an international sport, and global competitiveness has advanced significantly in recent years (Perret, 2017).

Moreover, in recent years, there has been a growing trend of utilizing technology in the field of sports. Various technologies are used to monitor performance in both competition and training to realize competitive advantages (Halson, 2014). In wheelchair sports, wheelchair movement performance is critical to evaluate game performance and optimize training routines (van der Slikke, 2016); however, literature related to wheelchair sports is limited compared to that of other Olympic events, and the quantitative evaluation of wheelchair movement performance is insufficient (Perret, 2017). Thus, the goal of this study is to realize an affordable tracking system to obtain kinematic data of wheelchair sports to improve wheelchair movement performance assessment. However, the literature related to wheelchair sports is scarce, and the quantitative evaluation of wheelchair movement performance is insufficient (Perret, 2017). This study aims to realize a tracking system to obtain kinematic data of wheelchair sports in order to improve the level of movement performance assessment.

2 RELATED WORK

With the increasing competitiveness in wheelchair sports, the utilization of technologies has expanded (Grogan, 2012) (Laferrier, 2012). For example, to evaluate wheelchair movement performance, Inertial Measurement Unit (IMU) sensors attached to the wheelchair are used frequently due to their user-friendliness and low cost (Shepherd, 2018). IMU

[a] https://orcid.org/0000-0002-8513-0204

Aihara, S., Sakai, T. and Shionoya, A.
Development of Monocular Vision-Based Tracking Method for Wheelchair Sports.
DOI: 10.5220/0012203600003587
In *Proceedings of the 11th International Conference on Sport Sciences Research and Technology Support (icSPORTS 2023)*, pages 179-186
ISBN: 978-989-758-673-6; ISSN: 2184-3201

sensors register movement, speed, and angular velocity of the wheelchair player (Pansiot, 2011) (van Dijk, 2022). In addition, methods based on IMU sensors can measure wheelchair movement easily. However, wheelchair positions are not accurately defined with the help of only IMUs accurate, which is one of the problems (van der Slikke, 2017).

The local positioning system (LPS) using wireless technology is a method that can be used to track wheelchair positions in various sports, e.g., wheelchair rugby and basketball (Rhodes, 2014), as well as wheelchair tennis (Perrat, 2015). However, to the best of our knowledge, very few related studies have been reported.

Wireless LPSs measure the position and speed of a wheelchair player at high accuracy. They comprise many fixed base stations and mobile tags attached to the wheelchair players; thus, there are some problems, e.g., the need for installation of base stations at the venue and expensive equipment. In addition, LPSs cannot obtain wheelchair motion direction. In wheelchair sports, the chair-work skill is an important factor when evaluating performance; thus, wheelchair direction information must be an available (Mason, 2013). In addition, attaching IMU sensors or LPSs to the wheelchairs or players causes several issues, e.g., preventing movement. In addition, such devices are not always permitted during official competitions.

Video-based tracking systems have been reported in the reference. Such systems eliminate the need to attach devices to the wheelchairs or players, and, therefore can collect data about all athletes in the game. In field sports, e.g., soccer, player positions can be acquired using deep learning techniques and image analysis from multiple cameras placed around the field (Redwood, 2012) (Linke, 2020). These techniques are used in FIFA and Japanese professional league matches. However, construction work and expensive equipment are required; thus, the costs of such systems are high.

Therefore, less expensive methods have been proposed to obtain player positions using single camera images (Buric, 2019) (Zhang, 2020). In the wheelchair rugby context, a previous study reported the acquisition of player positions using single camera images. Here, the wheelchair player detection rate and the position accuracy were approximately 20% lower than in the case of soccer; thus, operator corrections were required (Sarro, 2008). To the best of our knowledge, no video-based tracking system that obtains wheelchair directions has been reported to date.

In this study, we developed a video-based tracking method for wheelchair sports. Figure 1 shows an outline of the proposed method. To realize a simple and low-cost system, only a single camera is employed in the proposed method. The proposed system output both the wheelchair players' positions and the wheelchair motion directions. Thus, the proposed method represents a novel technique to acquire the variety data that has been difficult to get using other systems.

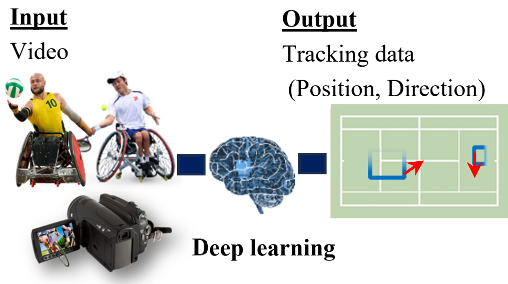

Figure 1: Outline of the proposed method.

3 PROPOSED METHOD

Figure 2 shows the steps followed to develop the proposed method. In the following, we first describe the development of the detection model, including dataset creation and the design of the model). We then describe the development of the tracking model, including the tracking model design, camera calibration, and the calculations used to acquire the position and direction information.

1. Detection model
   - Dataset creation
   - Detection model design
2. Tracking model
   - Tracking model design
   - Camara calibration
   - Position and direction calculation

Figure 2: Steps to develop the proposed method.

3.1 Detection Model

Here, we describe the development of the model used to detect a wheelchair player in the acquired images. Previous studies have reported monocular camera-based tracking methods that use the YOLO method (Redmon, 2016) to detect a bounding box (i.e., a square region around each player in the image) (Buric, 2019) (Zhang, 2020). However, these previous studies were limited to able-bodied athletes. When applied to a wheelchair player, bounding boxes were only detected for the player, i.e., the wheelchair was not included. Thus, the positions of the wheelchair players were identified as being above the ground, and correct positions could not be detected. Therefore, an effective model to accurately detect wheelchairs is required.

In previous studies (Buric, 2019) (Zhang, 2020), the center of the bounding box or the midpoint of the lower edge of the bounding box was used as the player position. In the current study, we developed a method to estimate the wheelchair structure by applying a human posture estimation model. Based on this model, the bottom points of each wheel are detected, and the midpoint of the bottom points of each wheel is calculated as the wheelchair player's position (shown in Figure 3). This method can detect the wheelchair player's positions using a monocular camera independent of the camera positions, wheelchair directions, and wheelchair player's posture.

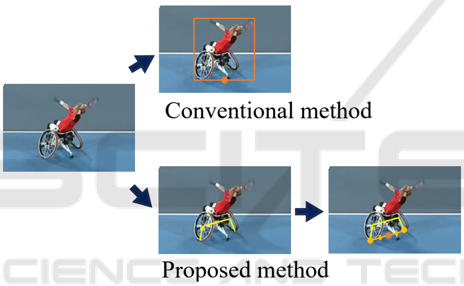

Figure 3: Wheelchair and player detection by the proposed model.

3.1.1 Dataset Creation

To the best of our knowledge, methods to estimate wheelchair structure from a camera image have not been reported. Thus, we developed a model to estimate the wheelchair structure. A marker-less posture estimation technique that uses a camera image requires a large dataset to optimize a large number of parameters. Thus, developing a wheelchair structure model from scratch required a large dataset of images with corresponding wheelchair key point coordinates. However, it is difficult to collect a large number of images of a specific category, e.g., wheelchair sports. In addition, it is difficult to construct a large dataset because this requires a lot of time and effort. Thus, in this study, we adapted a retraining method (Dai, 2015) that converts the human pose estimation model trained on the MS COCO library (Lin, 2014), which is a large human pose dataset. A new wheelchair structure estimation model can be created even with a small wheelchair sports dataset. In this study, we constructed a wheelchair sports dataset containing the feature points of wheelchair players and their wheelchairs to be used in the retraining method.

Figure 4 shows the key point coordinates. Here, the key points included the facial parts and the upper body joint points, in reference to the MS COCO data used in the pretraining process. The wheelchair key points were the centers of the left and right wheels and the bottoms of each wheel, which are common to all wheelchairs and can be used to capture the structure of the wheelchair effectively. In this study, a total of 17 key points (i.e., nose, left eye, right eye, left ear, right ear, left shoulder, right shoulder, left elbow, right elbow, left wrist, right wrist, left hip, right hip, center of left wheel, center of right wheel, bottom of left wheel, and bottom of right wheel) were defined in the wheelchair sports dataset. The knees and ankles defined in the MS COCO dataset were replaced by the centers of each wheel and the bottoms of each wheel in the wheelchair sports dataset. These changes were implemented to facilitate efficient fine tuning of the parameters. The process used to construct the wheelchair sports dataset is described as follows.

Step 1. Automatically collect (royalty-free) wheelchair sports images from the Internet.
Step 2. Normalize the image resolution (to 640 × 380 dpi).
Step 3. Mask people in the images who were not related to the wheelchair players (e.g., referees and spectators).
Step 4. Annotate the 17 key points.

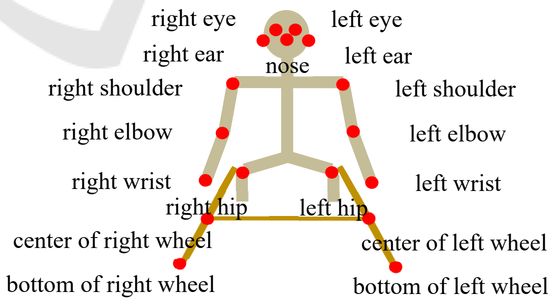

Figure 4: Key point coordinates.

In total, the wheelchair sports dataset contained approximately 2300 images with approximately 6000 subjects. The dataset included images of wheelchair basketball, rugby, tennis, badminton, and track and field. The images were collected from various wheelchair players in terms of gender and ethnicity. The pixel size and posture of the wheelchair players in the images also differed, and some of the images included overlapping wheelchair player images. The key point coordinates on the images were annotated manually by sports biomechanics experts. In addition, the data were divided randomly into training and test sets at a ratio of 7:3.

3.1.2 Detection Model Design

We adapted a human posture estimation model pretrained on the large-scale MS COCO dataset (Lin, 2014), and we retrained it on the acquired wheelchair sports dataset. As the foundation model, we used the Mask R-CNN (He, 2017), which is a widely used, flexible, and generic framework for human posture estimation methods. The architecture of the foundation model is shown in Figure 5. As shown, this model comprises three networks, i.e., the backbone network to extract features from the RGB images, the region proposal network to detect the regions of players and wheelchairs, and the key point branch to extract the key point coordinates of the players and wheelchairs. Thus, fine tuning the parameters of the three networks was required in this study.

First, the initial parameter weights were obtained by pretraining the algorithms on the MS COCO dataset, which contains posture information for approximately 150,000 humans in approximately 60,000 images. Then, the training data from the acquired wheelchair sports dataset were used for fine tuning. The optimization function was the Adam optimizer (Kingma, 2014) with a learning rate of 0.01. Here, 30% one of the training data was used as validation data, and the parameter weights with the minimized loss in the validation data were selected. The source code was implemented in Open CV, Python, and PyTorch, and training was performed using an NVIDIA Tesla V100 GPU on Google Colaboratory.

As a result, a new posture estimation model was developed that outputs the key point coordinates of the wheelchair player. Note that only the coordinates

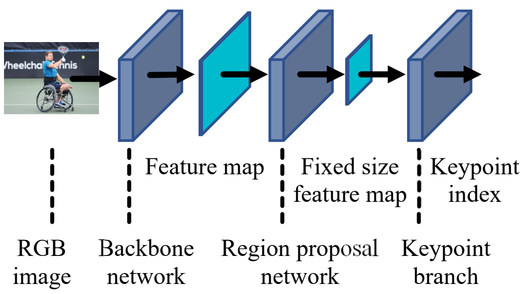

Figure 5: Architecture of the fundamental model.

of the bottoms of each wheel were used in the tracking method.

3.2 Tracking Model

3.2.1 Tracking Model Design

In Section 3.1, we described the model used to detect the key point coordinates (i.e., the coordinates of the bottoms of each wheel) in the images. However, this model was insufficient for the overall task due to occlusion caused by players overlapping images or motion blur caused by quick movements. Thus, we employed a model that tracks the bottoms of each wheel of the same player's and corrects the missing frames by linking a series of detection results between video frames. Here, we used the Byte Track method (Zhang, 2022) to track multiple objects. Byte Track links the detection results between frames by predicting the frame-to-frame changes in key point regions using a Kalman filter. This simple algorithm provides high stability, high speed, and high accuracy. The algorithm can stably track the bottom of each wheel of the same athlete throughout the entire video.

3.2.2 Camera Calibration

In this section, we describe the process of converting the key point coordinates (i.e., the coordinates of the bottom of each wheel) in the detected and tracked video image into a global coordinate (i.e., the position in the field).

We found that there was not possible to measure the coordinates of the calibration points in real time, and it was impossible to enter the target space for tracking. Thus, it was necessary to calculate the camera parameters from the feature points in the game or practice fields captured by the camera.

The game and practice fields have feature points whose length and size were specified by the International Sports Federation's regulation. Here, the camera parameters were calculated based on these feature points. Figure 6 shows an image illustrating the calculation of camera parameters on a wheelchair tennis court. The calibrator (i.e., the court model) was created using the court information specified by the regulations. Using this court model, the corresponding points of the global coordinates were mapped to pixel coordinates (at least four points) in the image.

The external camera parameters indicating the camera position and orientation in three-dimensional space were calculated using the Levenberg–Marquardt algorithm (Moré, 1978). The internal camera parameters indicating the focal distance of the camera were calculated using the hill climbing method (Goldfeld, 1966). Using these camera parameters, the image-based coordinate points were converted to global coordinate points, and by exchanging the court model, it is possible to adapt the algorithm to other sports.

The camera parameters were calculated using a single frame in the video; thus, the method did not support cases where the external camera parameters changed in the same video (e.g., camera pan, tilt, zoom, and position shift).

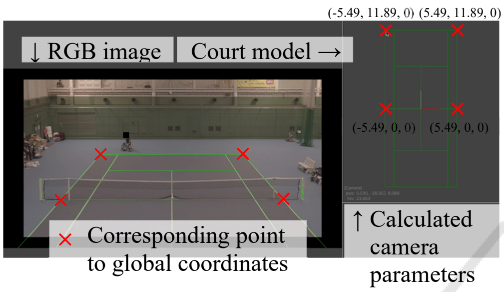

Figure 6: Camera calibration.

3.2.3 Positions and Directions Calculation

The two-dimensional (2D) position of each wheelchair player on the field and the corresponding wheelchair directions were calculated from the global coordinates of the key points (i.e., the bottom of each wheel). Figure 7 shows the coordinate frames in the case of wheelchair tennis. The 2D position of the wheelchair player (x_c, y_c) was calculated as the midpoint of the bottom of each wheel.

$$(x_c, y_c) = ((x_r + x_l)/2\,, (y_r + y_l)/2) \tag{1}$$

The wheelchair direction angle θ was calculated as follows. This represented the sagittal plane angle of the wheelchair.

$$\theta = arctan(-(x_r - x_l)/(y_r - y_l)) \tag{2}$$

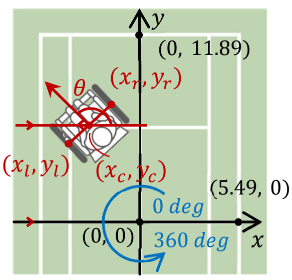

Figure 7: Example of global coordinate and direction in wheelchair tennis.

4 RESULTS

4.1 Accuracy of Pose Estimation Model

We evaluated the accuracy of the detection models developed (Section 3.1) on the wheelchair sports dataset using the test data, which were not used for training. Table 1 shows the error of each model. Here, the unit is pixels. The results for the person's posture are the average of the errors for each key point (i.e., eyes, nose, ears, shoulders, elbows, wrists, and hips), the results for the wheelchair structure are the average of the errors for each key point (i.e., the centers and bottoms of each wheel), and the results for the person and wheelchair structure are the average of the errors for all key points. For the human key point, the mean absolute error (MAE) was 4.43 pixels. The widely used methods for human posture estimation, Mask R-CNN (He, 2017) and Open Pose (Cao, 2021), were 4.68 and 4.51 pixels. Thus, the MAE value obtained by the proposed method was greater than that of the existing methods. The proposed method improved the estimation error by more than 1.7% compared to the existing methods (He, 2017) (Cao, 2021). For the wheelchair key points, the MAE was 6.22 pixels. These results confirm that the proposed method can be applied to various types of wheelchair sports, scenes, and individuals, as shown in Figure 8.

Table 1: MAE between the estimated coordinates and manually annotated coordinates (unit: pixels).

	Human pose	Wheelchair pose	Human and wheelchair pose
Mask R-CNN (He, 2017)	4.68	-	-
OpenPose (Cao, 2021)	4.51	-	-
Proposed method	4.43	6.22	4.94

Figure 8: Examples of estimation results by the proposed method (cropped to focus on wheelchair and humans).

4.2 Accuracy of Tracking Model

4.2.1 Data Collection

To evaluate the accuracy of the tracking model, an experiment was conducted during the wheelchair tennis matches. Six elite Japanese tennis players participated in the study. The matches were held on an indoor tennis court, and the players used the same wheelchairs they typically use in competitions. The players were divided into three groups and played one game (singles match). The players were requested to play with the same intensity as in international competitions. Two cameras (Pocket Cinema Camera 4K by Blackmagic Design Pty. Ltd., Port Melbourne, Australia) were placed at each corner of the tennis court. The height of the camera position was approximately 6 m. Each camera monitored half of the court. The resolution was 4K, and the frame rate was 60 fps. Present study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Ethics Committee of Nagaoka University of Technology.

4.2.2 Accuracy of Player Detection

The tracking data of the wheelchair players were output from the video images using the proposed method. Figure 9 shows an example of the tracking result. The player's trajectory was overlaid on the input image. Table 2 shows the detection rate of each wheelchair player. As can be seen, the proposed method was able to detect the wheelchair players in all frames.

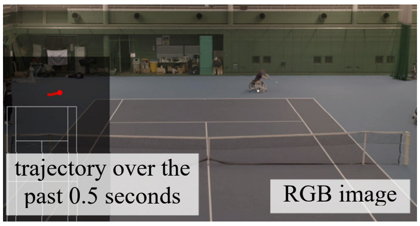

Figure 9: Image of the tracking results.

Table 2: Detection success rate of wheelchair player using the proposed method.

	Total data [s]	Detection data [s]	Detection rate [%]
Player 1	2650	2650	100
Player 2	2650	2650	100
Player 3	2300	2300	100
Player 4	2300	2300	100
Player 5	2900	2900	100
Player 6	2900	2900	100
All	15700	15700	100

4.2.3 Accuracy of Player Position

The positions of the wheelchair players were calculated using the proposed method. The videos were also digitized manually as reference values for validation. Here, for each player, 120 frames were selected randomly, and the bottoms of each wheel were digitized manually. The midpoint of each wheel was taken as the true value, and the coordinate transformation by the camera calibration was the same the proposed method. Table 3 shows the position determination errors of the proposed method (coordinate frames - according to Figure 7). The MAE in the horizontal direction (x) was 0.03 m, in the depth direction (y) was 0.10 m, and the root mean square error (RMSE) was 0.11 m. The values of "All" in Table 3 were calculated from all data of all players.

Table 3: Position determination errors of our method.

	MAE x [m]	MAE y [m]	RMSE [m]
Player 1	0.03	0.09	0.10
Player 2	0.03	0.09	0.10
Player 3	0.03	0.14	0.15
Player 4	0.02	0.09	0.09
Player 5	0.03	0.09	0.10
Player 6	0.03	0.12	0.12
All	0.03	0.10	0.11

4.2.4 Accuracy of Wheelchair Direction

The wheelchair motion directions of the players were calculated using the proposed method. As in the evaluation of the positional errors, here, the true values of the wheelchair directions were calculated by digitized manual data. Table 4 shows the wheelchair directions errors of the proposed model, and the coordinate system is shown in Figure 7. As can be seen, the mean ± SD was −2.23 ± 8.57 degrees, and the MAE was 6.78 degrees. The values of "All" in Table 4 were calculated from all data of all players.

Table 4: Wheelchair direction error of our method.

	Mean ± SD [deg]	MAE [deg]
Player 1	−3.22 ± 8.65	7.82
Player 2	−2.48 ± 9.89	8.02
Player 3	−2.80 ± 8.71	7.13
Player 4	−0.40 ± 7.98	6.34
Player 5	−0.73 ± 8.32	6.61
Player 6	−3.75 ± 7.05	6.46
All	−2.23 ± 8.57	6.78

5 DISCUSSIONS

We found that the proposed method estimated the wheelchair structure with the same accuracy as existing human posture estimation models, as shown in Table 1. Note that the constructed wheelchair sports dataset includes images of wheelchair basketball, rugby, tennis, badminton, and track and field; thus, the proposed method provides tracking data for a wide variety of wheelchair sports. In addition to wheelchair structure, the proposed model estimates the player’s upper body posture. In addition to measuring the wheelchair player positions, it is also possible to analyze the upper body movements. For example, the proposed method could be used to evaluate the wheelchair rowing motion. Also, it would be possible to analyze the relationship between the upper body usage and chair work skill using the proposed method.

As shown in Table 2, the detection success accuracy for wheelchair players were 100%. In a previous study (Sarro, 2008), the detection successrate of wheelchair rugby players using a video camera was approximately 74%, and that of wheelchair soccer players was approximately 94%. The proposed method demonstrates higher accuracy than previous study, and it achieved the accuracy required for use in sports.

The RMSE of the positions determinated by the proposed method was 0.11 m, as shown in Table 3. When using an LPS in wheelchair sports, the MAE was 0.19–0.32 m in wheelchair rugby and wheelchair basketball, respectively (Rhodes, 2014), and the MAE was 0.37 m for wheelchair tennis (Perrat, 2015). The results obtained for the proposed method indicate that it outperforms these existing methods in terms of accuracy.

In this study, the position accuracy was evaluated in wheelchair tennis. The proposed method can be applied to other wheelchair sports by exchanging the court model for camera calibration. The proposed method is an innovative tracking system that does not require base stations or devices attached to the players, and it can realize high position detection accuracy using only a single camera.

The MAE of the wheelchair directions tracked by the proposed method was 6.78 degrees, as shown in Table 4. Methods based on a single IMU sensor are widely used to measure wheelchair directions. For example, previous studies reported 8.1 degrees (van Dijk, 2022) and 11.0 degrees (Rupf, 2021). Thus, the proposed method outperforms methods based on a single IMU sensor. In addition, to the best of our knowledge, the proposed method is the first based on a monocular camera. Thus, the proposed method provides a simplified novel tool to obtain kinematic data for wheelchair sports.

The proposed method provides the movement information of wheelchair players using a single camera placed at the side of the field or near audience seats. Thus, it is useful for training load management and evaluating on-court performance. In addition, for competitive sports, the proposed method can be used to acquire kinematic data of opponents to improve the analysis of tactics.

The proposed method can be applied to the analysis of past legendary players and to compare the past and current performance of the same player, even if it is not possible to acquire new data using the tracking system. We believe that our findings contribute to the quantitative performance evaluation of wheelchair athletes.

Finally, we describe the limitations observed in this study. We found that the proposed method exhibits a larger error in the depth direction than in the horizontal direction due to the single camera (Table 3). In addition, the position error increases when the number of pixels per wheelchair player decreases. Thus, with the proposed method, it is necessary to consider the image acquisition conditions. In this study, the evaluation was conducted for singles wheelchair tennis; thus, the accuracy may decrease according to the overlap of wheelchair players. Therefore, in the future, we plan to evaluate the proposed method when multiple players are present on the same field.

6 CONCLUSIONS

In this study, we developed a tracking method to measure the kinematic data of wheelchair sports using a monocular camera. With the proposed method, the RMSE of the wheelchair player position was 0.11 m, and the MAE of the wheelchair direction was 6.78 degrees. In addition, the proposed method achieved higher accuracy than existing tracking methods, e.g., the LPS and IMU sensor–based methods. The proposed method provides a simple tool to obtain kinematic data in wheelchair sports, which have not been collected previously. This research contributes to the quantitative performance evaluation of wheelchair athletes.

ACKNOWLEDGEMENTS

This work was supported by the "Functional Development Project for Resilient Athlete Support" of Japan Sports Agency.

REFERENCES

Buric, M., Ivasic-Kos, M., & Pobar, M. (2019). Player tracking in sports videos. *Proceedings of the International Conference on Cloud Computing Technology and Science*, 334–340.

Cao, Z., Hidalgo, G., Simon, T., Wei, S. E., & Sheikh, Y. (2021). OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(01), 172–186.

Dai, A. M., & Le, Q. V. (2015). Semi-supervised Sequence Learning. *Advances in Neural Information Processing Systems*, 28.

Goldfeld, S. M., Quandt, R. E., & Trotter, H. F. (1966). Maximization by Quadratic Hill-Climbing. *Econometrica*, 34(3), 541.

Grogan, A. (2012). Paralympic technology. *Engineering and Technology*, 7(8), 28–31.

Halson, S. L. (2014). Monitoring Training Load to Understand Fatigue in Athletes. *Sports Medicine*, 44(2), 139–147.

He, K., Gkioxari, G., Dollar, P., & Girshick, R. (2017). Mask R-CNN. In *Proceedings of the IEEE International Conf. on Computer Vision*, 2961–2969.

Kingma, D. P., & Ba, J. L. (2014). Adam: A Method for Stochastic Optimization. 3rd International Conference on Learning Representations.

Laferrier, J. Z., Rice, I., Pearlman, J., Sporner, M. L., Cooper, R., Liu, T., & Cooper, R. A. (2012). Technology to improve sports performance in wheelchair sports. *Sports Technology*, *5*(1–2), 4–19.

Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., & Zitnick, C. L. (2014). Microsoft COCO: Common objects in context. *European Conf. on Computer Vision*, 8693, 740–755.

Linke, D., Link, D., & Lames, M. (2020). Football-specific validity of TRACAB's optical video tracking systems. *PLOS ONE*, 15(3), e0230179.

Mason, B. S., Porcellato, L., Van Der Woude, L. H. V., & Goosey-Tolfrey, V. L. (2010). A qualitative examination of wheelchair configuration for optimal mobility performance in wheelchair sports: a pilot study. *Journal of Rehabilitation Medicine*, 42(2), 141–149.

Moré, J. J. (1978). The Levenberg-Marquardt algorithm: Implementation and theory. *Numerical Analysis*, 105–116.

Pansiot, J., Zhang, Z., Lo, B., & Yang, G. Z. (2011). WISDOM: wheelchair inertial sensors for displacement and orientation monitoring. *Measurement Science and Technology*, 22(10), 105801.

Perrat, B., Smith, M. J., Mason, B. S., Rhodes, J. M., & Goosey-Tolfrey, V. L. (2015). Quality assessment of an Ultra-Wide Band positioning system for indoor wheelchair court sports. *Proceedings of the Institution of Mechanical Engineers, Part P: Journal of Sports Engineering and Technology*, 229(2), 81–91.

Perret, C. (2017). Elite-adapted wheelchair sports performance: a systematic review. *Disability and Rehabilitation*, 39(2), 164–172.

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 779–788.

Redwood-Brown, A., Cranton, W., & Sunderland, C. (2012). Validation of a real-time video analysis system for soccer. *International Journal of Sports Medicine*, 33(8), 635–640.

Rhodes, J., Mason, B., Perrat, B., Smith, M., & Goosey-Tolfrey, V. (2014). The validity and reliability of a novel indoor player tracking system for use within wheelchair court sports. *Journal of Sports Sciences*, 32(17), 1639–1647.

Rupf, R., Tsai, M. C., Thomas, S. G., & Klimstra, M. (2021). Original article: Validity of measuring wheelchair kinematics using one inertial measurement unit during commonly used testing protocols in elite wheelchair court sports. *Journal of Biomechanics*, 127.

Sarro, K. J., Misuta, M. S., Burkett, B., Malone, L. A., & Barros, R. M. L. (2010). Tracking of wheelchair rugby players in the 2008 Demolition Derby final. *Journal of Sports Sciences*, 28(2), 193–200.

Shepherd, J. B., James, D. A., Espinosa, H. G., Thiel, D. V., & Rowlands, D. D. (2018). A Literature Review Informing an Operational Guideline for Inertial Sensor Propulsion Measurement in Wheelchair Court Sports. *Sports 2018*, 6(2), 34.

an der Slikke, R. M. A., Berger, M. A. M., Bregman, D. J. J., & Veeger, H. E. J. (2016). From big data to rich data: The key features of athlete wheelchair mobility performance. *Journal of Biomechanics*, 49(14), 3340–3346.

van der Slikke, R. M. A., Mason, B. S., Berger, M. A. M., & Goosey-Tolfrey, V. L. (2017). Speed profiles in wheelchair court sports; comparison of two methods for measuring wheelchair mobility performance. *Journal of Biomechanics*, 65, 221–225.

van Dijk, M. P., van der Slikke, R. M. A., Rupf, R., Hoozemans, M. J. M., Berger, M. A. M., & Veeger, D. J. H. E. J. (2022). Obtaining wheelchair kinematics with one sensor only? The trade-off between number of inertial sensors and accuracy for measuring wheelchair mobility performance in sports. *Journal of Biomechanics*, 130, 110879.

Zhang, Y., Chen, Z., & Wei, B. (2020). A Sport Athlete Object Tracking Based on Deep Sort and Yolo V4 in Case of Camera Movement. *2020 IEEE 6th International Conf. on Computer and Communications*, 1312–1316.

Zhang, Y., Sun, P., Jiang, Y., Yu, D., Weng, F., Yuan, Z., Luo, P., Liu, W., & Wang, X. (2022). ByteTrack: Multi-object Tracking by Associating Every Detection Box. *European Conf. on Computer Vision*, 13682, 1–21.