PADELVIC: MULTICAMERA VIDEOS AND MOTION CAPTURE DATA IN PADEL MATCHES

PADELVIC: VIDEOS MULTICÁMARA Y DATOS DE CAPTURA DE MOVIMIENTO EN PARTIDOS DE PÁDEL

MOHAMMADREZA JAVADIHA
Universitat Politècnica de Catalunya.
Orcid: 0000-0002-4867-1132

CARLOS ANDUJAR GRAN
Universitat Politècnica de Catalunya.
Orcid: 0000-0002-8480-4713

MICHELE CALVANESE
Università di Bologna.
Orcid: 0009-0005-4103-0147

ENRIQUE LACASA
Universitat Politècnica de Catalunya.
Orcid: 0000-0002-4171-9113

JORDI MOYÉS
Universitat Politècnica de Catalunya.
Orcid: 0009-0002-4076-5676

JOSÉ LUIS PONTÓN
Universitat Politècnica de Catalunya.
Orcid: 0000-0001-6576-4528

ANTONIO SUSÍN
Universitat Politècnica de Catalunya.
Orcid: 0000-0002-0874-2784

JIABO WANG
Universitat Politècnica de Catalunya.
Orcid: 0009-0002-3243-3110

Autor de correspondencia: Carlos Andujar Gran, Universidad Politénica de Catalunya. Campus Nord UPC, Omega 239, Jordi Girona, 1-3, 08034, Barcelona. España. andujar@cs.upc.edu

ABSTRACT

Recent advances in computer vision and deep learning techniques have opened new possibilities regarding the automatic labeling of sport videos. However, an essential requirement for supervised techniques is the availability of accurately labeled training datasets. In this paper we present PadelVic, an annotated dataset of an amateur padel match which consists of multi-view video streams, estimated positional data for all four players within the court (and for one of the players, accurate motion capture data of his body pose), as well as synthetic videos specifically designed to serve as training sets for neural networks estimating positional data from videos. For the recorded data, player positions were estimated by applying a state-of-the-art pose estimation technique to one of the videos, which yields a relatively small positional error (M=16 cm, SD=13 cm). For one of the players, we used a motion capture system providing the orientation of the body parts with an accuracy of 1.5º RMS. The highest accuracy though comes from our synthetic dataset, which provides ground-truth positional and pose data of virtual players animated with the motion capture data. As an example application of the synthetic dataset, we present a system for a more accurate prediction of the center-of-mass of the players projected onto the court plane, from a single-view video of the match. We also discuss how to exploit per-frame positional data of the players for tasks such as synergy analysis, collective tactical analysis, and player profile generation.

Keywords: Computer vision, pose estimation, player tracking, sport analytics.

RESUMEN

Los avances en visión por computador y el aprendizaje profundo han abierto nuevas posibilidades en cuanto a la etiquetación automática de videos deportivos. Sin embargo, un requisito esencial para las técnicas de aprendizaje supervisado es la disponibilidad de datos de entrenamiento con anotaciones precisas. En este artículo presentamos PadelVic, un conjunto de datos anotado de un partido amateur de pádel que incluye múltiples secuencias de video desde diferentes ángulos, estimaciones de la posición en pista de los cuatro jugadores (y datos de captura de movimiento de uno de los jugadores), así como videos sintéticos diseñados específicamente para ser utilizados como entrenamiento para redes neuronales que estimen la posición a partir de videos. La posición de los jugadores se obtuvo aplicando un método de estimación de pose a uno de los videos, el cual proporciona un error posicional (M=16 cm, SD=13 cm) razonable en comparación con el tamaño de la pista. Para uno de los jugadores, utilizamos un sistema de captura de movimiento que proporciona la orientación de las partes del cuerpo con una precisión de 1.5º RMS. Sin embargo, el mayor grado de exactitud proviene de nuestros datos sintéticos, que proporcionan la posición y

orientación exacta de jugadores virtuales que imitan los datos del sistema de captura de movimiento. Como ejemplo de aplicación de los datos sintéticos, presentamos un sistema para predecir con mayor precisión el centro de masa de los jugadores proyectado sobre el plano de la pista, a partir de un solo video del partido. También discutimos cómo aprovechar los datos posicionales de los jugadores para tareas como análisis de sinergia, análisis táctico colectivo, y generación de perfiles de jugadores.

Palabras clave: visión por computador, estimación de la pose, seguimiento de jugadores, análisis deportivo.

Introduction

In the last few years, the increased demand of padel from both players and spectators has also stimulated an increase of the scientific production about this interceptive sport. In particular, many recent papers deal with the analysis of aspects influencing performance (Courel-Ibáñez et al., 2015; Sánchez-Alcaraz, Cánovas et al., 2022; Martín-Miguel et al., 2023). The analysis of technical actions (Escudero-Tena et al., 2022; Conde-Ripoll, et al., 2021), tactical situations (Ramon-Llín et al., 2021), or external load parameters (Sánchez-Alcaraz et al., 2021) has also facilitated the understanding of the game. Consequently, coaches and trainers are increasingly interested in incorporating these elements into their training plans.

Most of these analyses require, or could benefit from, positional data about the players and the ball. Indeed, the ball-flight information and the upper body of the opponents (head, shoulders, trunk, and the region of arm–hand–racket) were the two most relevant visual locations for experienced padel players because they made larger number of fixations and longer fixation times at these sources of information when returned three representative situations performed by their counterparts (trays/smashes, serves, and volleys) on a padel court (Espino et al., 2023). Although other types of information (e.g., the contextual information as opponents´ tendencies, weakness of the counterparts, etc.) is certainly needed for an appropriate analysis of the game, getting per-frame positional data from a video is much more involved task than other important but less frequent events such as the type of stroke of the player returning the ball.

Since annotating manually the players' positions and poses in every single frame of a video is a very involved task, requiring a substantial amount of effort, the automatic annotation of sport matches opens new opportunities regarding the amount of padel matches that can be subject to analysis.

A variety of sensors, cameras, and advanced algorithms allow us to track and measure different performance metrics such as player speed, acceleration, traversed distances, movement patterns, heart rate, calorie consumption, stress tolerance, and pressure, at different accuracy levels (Galeano et al. 2022).

The semi-automatic analysis of sport videos often involves three main steps. The first step involves player detection, identification and tracking, which often requires postprocessing filters to remove false positives. The second step is to identify and classify meaningful events of the match (such as shots, ball bounces, and scoring units such as points, games, and sets). Finally, after tracking data has been structured into meaningful units (e.g., rallies, games, sets) it can be visualized and analyzed. This analysis can be facilitated by padel-specific query languages to retrieve and analyze specific in-game situations from a match (Javadiha et al., 2022).

Computer vision techniques are already used in some racquet sport for position detection (Su et al., 2018), ball detection (Voeikov et al., 2020), shot detection (Horie et al., 2019), as well as spatio-temporal and visual analysis for tactics (Chu et al., 2021; Wu et al., 2021).

Although traditional object-detection techniques require little or no training data, current state-of-the-art supervised methods rely on deep learning and convolutional neural networks. An essential requirement for these techniques is the availability of accurately labeled training sets. Although many training sets have been released for general sports (Andriluka et al., 2014; 2018), to the best of our knowledge no training dataset specific to padel has been published.

In this paper we present an annotated dataset of an amateur padel match. The dataset includes videos from different camera angles, estimated court-space positions for the four players, accurate motion capture data of one of the players (acquired with a professional MoCap system), as well as synthetic videos with ground-truth positions and poses, specifically designed to serve as training sets for deep networks predicting positional data.

The automatic acquisition of such positional data opens up vast horizons for data analysis. Poses allow us to analyze the frequency of certain technical

actions, the potential diversity of players in competition (or in specific training situations), their evolution, as well as how their differences are influenced by the season, indoor/outdoor tournaments, and other factors. Exploratory behavior in different temporal sequences can be analyzed through methods like Dynamic Overlap and Principal Components Analysis (Hristovski et al., 2011). Furthermore, the acquisition of positional data would allow further understanding for synergy analysis (Passos et al., 2020), collective tactical principle analysis (Canton, 2022), player profile generation (Buldú et al., 2021), attractor identification (Galeano et al., 2022), and external load estimation in terms of distances covered and speeds. Depending on the researcher's interest, correlating some of these variables with rally outcomes and effectiveness metrics can provide valuable information for analysts, coaches and players. In any case, since the technique supports tactics while also being subordinate to tactics (Bonnet, 1986), positional data alone is not enough and it must be used together with other information such as the tactical intention of the players.

Method

Objectives

The overall goal of this project is to provide a public dataset for padel that can be used for different computer vision tasks in the specific context of padel matches. In particular, we aim to use the dataset as a technical solution for (a) testing the performance of computer vision algorithms in tasks such as video-based player tracking, pose estimation, ball tracking and shot recognition, and (b) training supervised algorithms that estimate diverse positional and/or pose-related data, such as the position and orientation of the padel racket with respect to the player, or the court-space position of the players.

The performance of player detection and pose estimation techniques largely depends on the number of cameras (e.g., in general, the larger, the better), their orientation (e.g., zenithal views allow for accurate court-space positioning of the players; court-level views improve pose estimation accuracy) and other factors such as video frame rate and image resolution. Since we strive to use the dataset in inexpensive setups affordable for most padel clubs, we focus on single-view applications (although the dataset includes multiple views to serve as test data).

On the other hand, we aim to provide camera views close to the de-facto standard in professional video streaming, which minimizes the impact of the occlusion of the mesh panel and structural elements. In this setting (Figure 1),

the camera angle causes the mesh panel to span the region from the bottom part of the net to near the opposite service line. Ideally, the camera should be placed at about 7.6 m above the floor, and 15.5 m far behind from the glass panels, but cameras above 5 m high already provide a reasonably similar view.

Figure 1. *De-facto standard for camera placement in padel broadcasting*
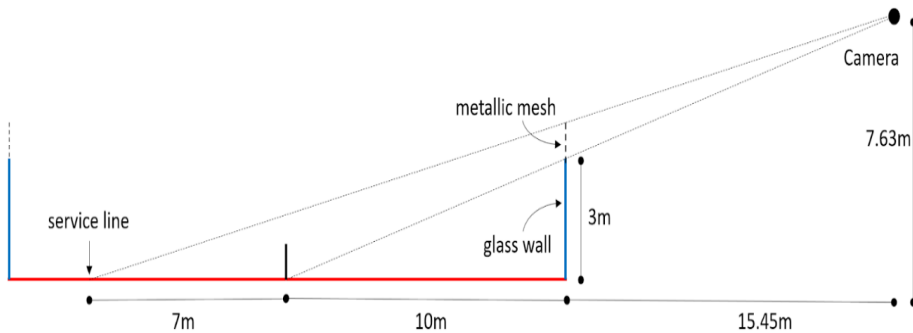
The above goals require that the dataset is enriched with annotations that accurately represent the motion of at least one player. For this reason, we used a professional, suit-based motion capture system to track the pose and position of one of the players (Figure 3-left). Finally, we used an open-source 3D modeling application (Blender) to animate synthetic characters with the captured motion data, which allows us to generate highly-accurate, annotated synthetic videos from arbitrary views.

After visiting several padel clubs with the aim of recording the padel match, Aurial Padel Vic was chosen since it allowed us to record the match from a position close to the de-facto standard. More precisely, Aurial Padel Vic has a two-story structure with an extendable ceiling panel at the top, which allowed us to place one of the cameras above it and thus reach the desired camera height and angle. Furthermore, the chosen court has a considerable free space around which largely facilitated recording the video without encountering significant hindrances, and even recording the match from other points of view with the help of secondary cameras. Another noteworthy aspect is that the court carpet is blue, which enhances the color contrast with the yellow ball. The match was recorded in winter (March 1st, 2023) to prevent player's sweat from affecting the sensors of the motion capture suit.

Captured datasets

We recorded four amateur players, all of them with more than 10 years of experience playing padel regularly.

For the *main camera view*, we used a Panasonic AG-UX180 4K professional camcorder. The camera was placed at about 6 m high from the court plane, and 10 m far away from the back wall, thus getting a camera angle close to the de-facto standard (Figure 2). The camera zoom was adjusted to enclose the complete court and part of its surroundings. The match was recorded at 4K UHDTV (3840x2160), 50 fps. The output video was about 86 GB in MP4 format.

Figure 2. *Sample frames from the captured videos*

Note. The figure shows the main camera view (*top left*), the second camera view (*top right*) and the court-level smartphone views (*bottom*).

In order to get a video with a more vertical angle, the *second camera* was placed at a slightly higher position. We used a GoPro Hero9 Black on a tripod above a ceiling panel structure with a stepped ceiling. The video was recorded at 4K UHDTV, 60 fps. The output video was about 32 GB in MP4 format. This video shows a lamp hanging from the ceiling which occludes a part of the top-left court.

For *court-level recordings*, we used two tripod-mounted smartphone cameras at about 1.45 m high, 3.88 m away from the court, with 82 cm lateral

separation which emulates a stereo camera. Both were oriented in landscape mode. We used a Samsung Galaxy S22+ to record FHD (1920x1080), 60 fps video (H.264, MP4, 17GB) and an IPhone 13 Pro Max was also set to record FHD (1920x1080) video (MOV format, 8GB).

Additionally, we tracked one of the participants playing the game, using a professional MoCap system. He was an experienced but amateur player, 51 years old, 1.70 cm, playing for more than 10 years and about 4 days/week, not federated, and not involved in any official competitions. In particular, we used an XSens MVN Awinda suit with 18 inertial measurement units (17 body-worn + 1 prop), Figure 3. We accurately captured the player's body poses and tracked the orientation and location of the padel racket with respect to the player. The resulting motion was captured at a frame rate of 60 fps and represented using a 24-joints skeleton. Our dataset provides the captured motion data in BVH format (~183MB, ~74MB compressed).

Table 1 summarizes the captured datasets.

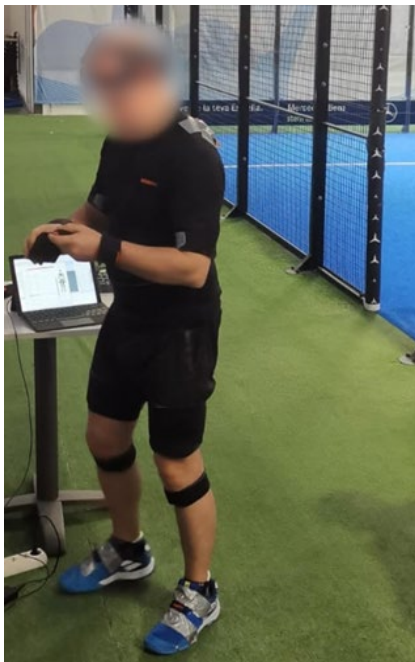**Figure 3.** *Player wearing XSens suit and sample captured frames*

Table 1. *Summary of captured datasets*

Dataset	Height	Device	Resolution	FPS	Format
Main camera	6 m	Panasonic AG-UX180 4K	3840x2160	50	MP4
Second camera	7.5 m	GoPro Hero9 Black	3840x2160	60	MP4
Court level 1	1.45 m	Samsumg Galaxy S22+	1920x1080	60	MP4
Court level 2	1.45 m	IPhone 13 Pro Max	1920x1080	60	MOV
Motion Capture	N/A	Xsens MVN Awinda	N/A	60	FBX/BVH

Generated datasets

We used a collection of scripts to generate new, annotated datasets from the captured data: synthetic videos showing an avatar that reproduces the motion and gestures of the tracked player, labeled with accurate positional and pose data, and a spreadsheet file describing the position and pose of the detected players, for each frame of active gameplay in the video.

Regarding positional data, object detection algorithms provide the object type indicated as a class number and its 2D location within the image (usually a bounding box enclosing the object). There are some pre-designed deep neural networks for detecting custom objects such as humans, rackets and balls. Such backbone architectures might be trained using different datasets such as COCO or PASCAL VOC. Apart from object identification and detection, pose estimation deals with the detection of major key joints in a human body, such as knees, ankles, hip, and shoulders, and then connecting these joints to reconstruct a body skeleton representing the pose. When presented a collection of frames or a video, one can tag every detected player based on their name or initial location in the scene, and associate the same labels to the same player on the following frames. This process is known as *player tracking*.

In our case study, we focus on the players, the ball and the racket. For human detection, we used HTC (Chen et al., 2019) to detect all players inside the court. More precisely, we used the htc_r101_fpn_20e backbone pretrained with the COCO2020 dataset (available in OpenMMLab). Then, we used the top-down HRNet approach (Chen et al., 2020) to obtain the pose of every single person inside the rectangles estimated by the HTC detector. Specifically,

we used the hrnet_w48_coco_256192 model (again available in OpenMMLab).

The hip and feet keypoints were used to estimate the player position in camera-space. Finally, we reversed the perspective deformation to obtain court-space coordinates of the players. For player tracking, we used a custom algorithm that matches detected persons with players based on the overlap of such detections across the video frames, using as overlap metric the IoU (Intersection Over Union). Figure 4 shows detected players and keypoints for a sample frame of our dataset.

For testing, we also applied a custom ball tracking method based on the TracknetV2 architecture (Huang et al., 2019) to detect the ball, followed by a RANSAC algorithm to extract the parabolic paths (Yan et al., 2006). The result for a sample frame is shown in Figure 5. Notice though that PadelVic dataset does not include explicitly ball tracking data.

Figure 4. *Detected poses on a sample frame of the recorded video*

Figure 5. *Ball tracking on a sample frame of the recorded video*
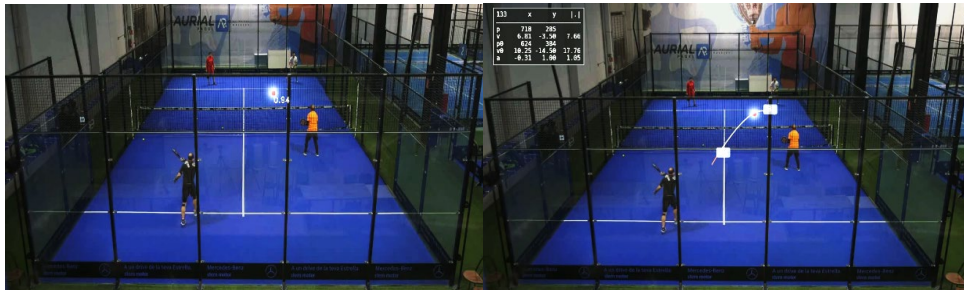
Note. The figure shows the detected ball (*left*) and the estimated trajectory in neighboring frames (*right*).

As for the synthetic videos, we created a Python script for the open source 3D modeling software *Blender* to automate the process of transferring the MoCap data to arbitrary avatars placed on a virtual padel court. The user can then select the desired camera angle, and render the animation to create synthetic annotated videos. The script also dumps labels to a CSV file whose columns include the frame number, the 3D position of the joints, as well as user-defined derived data such as the orthogonal projection of the feet and hip onto the court plane. The major benefit of this approach is that labels can refer to any measure that can be computed from the 3D MoCap data, and it can be represented either in 3D space coordinates (e.g. height of the dominant hand) or in image-space coordinates (by just projecting the 3D position onto the camera space). Furthermore, labels are pixel-accurate even when predicting keypoints hidden by clothes or other elements (in contrast to human-made annotations). As shown in Figure 6, the resulting Blender model can be rendered from arbitrary view points, enabling the training and testing of both single-camera and multi-camera approaches.

Figure 6. *Sample frames from a synthetic video included in the dataset*
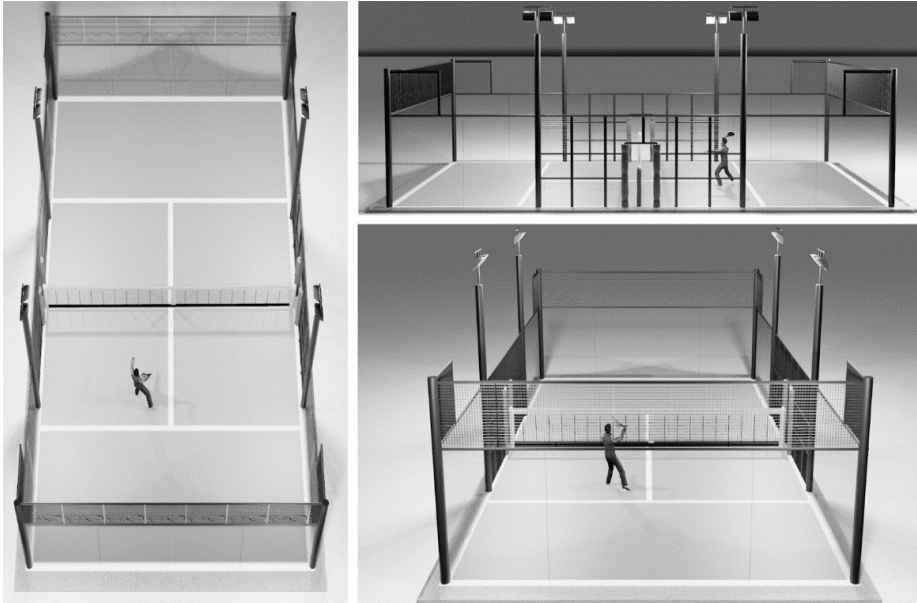
Synchronization issues

Achieving synchronicity between motion capture poses and video streams is nontrivial. Despite both types of data potentially being recorded at comparable frequencies, the absence of a unified timekeeper results in temporal drift (about 1 second per hour) and discrepancies in recording times.

The Xsens IMU-based MoCap suit enabled player motion capture without external sensors, thus avoiding occlusions and allowing capturing in an unrestricted space. Nonetheless, IMU sensors exhibit positional drift over extended periods due to the integration method employed for position computation. Therefore, positional information may not be robust enough for the synchronization task. To address this, we plan in future work to extract key postural features from the video data and employ a similarity algorithm to align this information with the motion-captured sequences. This approach could necessitate frame interpolation due to different capture frequencies, and its effectiveness could be affected by intermittent gaps in accurate pose data from the videos.

Sample application

We used the synthetic video and the labeled data (extracted automatically from the skeleton joints of the virtual player) to train a neural network that predicts the position of the players, as the projection of its hip onto the court. The projected position is more suitable for 2D spatial analysis of players' displacements. A convolutional neural network first estimates key points of the player in one or more neighboring frames, using the mentioned HTC method for detecting all players inside the court (Chen et al., 2019), and HRNet for obtaining the pose of these detected players (Chen et al., 2020). Then a simple fully-connected network takes as input the estimated poses and learns to predict the floor projection of the hip (Figure 7). This addresses the common problem of estimating the 2D position of the players in non-zenithal cameras: the camera angle combined with the inherent perspective deformation causes some players (e.g. when jumping) to be estimated in wrong positions (Figure 8).

The resulting neural network is able to predict the player position (projection of the player's hip onto the floor) with an average deviation of 10 pixels (on a FullHD video frame) in each direction. This error is reduced down to only 5 pixels if the network uses multiple frames before and after the current frame.

Figure 7. *Estimated poses and predicted court-space position of the player*

Note. The poses estimated from the synthetic video (*left*) allowed us to train a network to predict accurately the position of the player. Using one frame provides a good estimate of the court-projected position (*middle*), which becomes even more robust when using multiple frames (*right*). The red circle is the predicted position, whereas the green circle is the ground truth.

Figure 8. *Potential positional errors due to vertical displacements of the players*
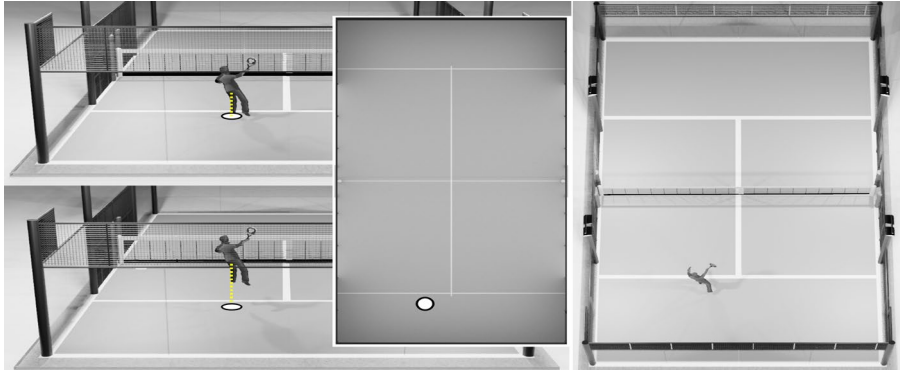
Note. Vertical displacements of the players (while running and jumping) severely influence the accuracy of the estimated player positions in court-space. The image on the left shows two players with different image-space positions, but identical court-space positions (middle and right images). Without proper handling, the position of jumping players could be estimated meters away from the correct one.

Strengths and limitations of padelvic

A first limitation is that the partipants' behaviors, including movements, technical actions and tactical decisions, were somewhat influenced by the capture conditions. Players were instructed to play as in usual amateur matches, although multiple factors may have affected their behavior and accuracy: the presence of multiple cameras, the frequent breaks due to calibration/tracking issues, and the cumbersomeness of the motion capture suit, especially the sensors in the wrist and in the racket. Although all players were recorded in the videos, only one was wearing the MoCap suit. This might have introduced some bias in the opponents, to favor the active participation of the tracked player. The inertial sensor mounted on the padel racket (16 g), together with the material to fix it, changed significantly the weight and balance of the racket. Finally, the tracker sensor for the hand in the XSens Awinda is attached to a glove which largely interferes with the natural motion of the hand, and its thickness prevents a proper grip of the racket. These aspects certainly influenced the choice, execution and accuracy of the technical actions of all players, especially of the tracked player.

The videos were recorded using commodity cameras which did not support external sync sources. After manual adjustment of the starting frame, the video streams we release only show an approximate synchronization, but not

per-frame sync. This would be a limitation if the multiple streams were to be used for multi-camera reconstruction. However, each video is still valid for testing single-camera techniques under different camera angles. The MoCap data included as part of PadelVic allows for the generation of synthetic videos which can be rendered from arbitrary camera angles, with perfect synchronization and ground-truth data. These videos can be used to train and test both single-camera and multi-camera approaches for the estimation of positional data from videos.

Conclusions and practical applications

We have presented a padel-specific collection of datasets targeted to train and test computer vision applications on padel videos. The dataset *(to be released upon acceptance, early preview version available at https://github.com/UPC-ViRVIG/PadelVic)* includes videos of an amateur match recorded from different camera angles, as well as MoCap data and synthetic videos that show player avatars mimicking the motion of the amateur player.

Example applications of the dataset include testing and training deep learning methods to predict derived positional data (e.g. tip of the racket, center of the racket, height of the ball with respect to the floor) from a single or multiple videos, requiring no external sensors (Lacasa et al., 2021).

Although PadelVic can be used to train and test ball detection and ball tracking algorithms, robust ball tracking is still a challenging problem due to occlusion and the presence of short path segments when the ball bounces off the walls. Accurate ball trajectories would allow trainers to analyze player activity patterns based on the trajectory of the ball and the position of players on the court. The work that Fernando Rivas has conducted with Carolina Marín in badminton in this regard (Gómez et al., 2020) should inspire the padel research community to generate knowledge along similar lines in padel.

As future work, we plan to explore how positional data about the ball and the players can be used for enhancing notational analysis. While so far we have compared some initial basic report proposals, the availability of accurate per-frame data opens new horizons for this field. We also plan to acquire, manually annotate and release additional datasets, with highly skilled professional padel players which perform movements more quicky and hit the ball harder than amateur players. A major challenge though is how to minimize the impact of the trackers (especially those mounted on the gloves and the

racket) on the players' motor actions.

Acknowledgments and funding

References

Andriluka, M., Pishchulin, L., Gehler, P., & Schiele, B. (2014). 2d human pose estimation: New benchmark and state of the art analysis. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 3686-3693). https://doi.org/10.1109/CVPR.2014.471

Andriluka, M., Iqbal, U., Insafutdinov, E., Pishchulin, L., Milan, A., Gall, J., & Schiele, B. (2018). Posetrack: A benchmark for human pose estimation and tracking. In Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 5167-5176).
https://doi.org/10.1109/CVPR.2018.00542

Cheng, B., Xiao, B., Wang, J., Shi, H., Huang, T. S., & Zhang, L. (2020). Higherhrnet: Scale-aware representation learning for bottom-up human pose estimation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 5386-5395).

Chen, K., Pang, J., Wang, J., Xiong, Y., Li, X., Sun, S., ... & Lin, D. (2019). Hybrid task cascade for instance segmentation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 4974-4983).

Chu, X., Xie, X., Ye, S., Lu, H., Xiao, H., Yuan, Z., ... & Wu, Y. (2021). TIVEE: Visual exploration and explanation of badminton tactics in immersive visualizations. *IEEE Transactions on Visualization and Computer Graphics*, *28*(1), 118-128. https://doi.org/10.1109/TVCG.2021.3114861

Conde-Ripoll, R., Llanos, M.B., García, J.M., & Sánchez-Alcaraz, B.J. (2021). Análisis de los golpes de pared en pádel profesional. Revista de Entrenamiento Deportivo, 35(2), 3-11.

Courel-Ibáñez, J., Sánchez-Alcaraz, B.J., & Cañas, J. (2015). Effectiveness at the net as a predictor of final match outcome in professional padel players. International Journal of Performance Analysis in Sport, 15(2), 632-640. https://doi.org/10.1080/24748668.2015.11868820

Escudero-Tena, A., Almonacid, B., Martínez, J., Martínez-Gallego, R., Sánchez-Alcaraz, B.J., & Muñoz, D. (2022). Analysis of finishing actions in men's and women's professional padel. International Journal of Sports

Science&Coaching. https://doi.org/10.1177/17479541221139970

Espino Palma, C.; Luis del Campo, V.; Muñoz Marín, D. (2023). Visual Behaviours of Expert Padel Athletes When Playing on Court: An In Situ Approach with a Portable Eye Tracker. Sensors 2023, 23, 1438. https://doi.org/10.3390/s23031438

Galeano, J., Gómez, M. Á., Rivas, F., & Buldú, J. M. (2022). Using Markov chains to identify players' performance in badminton. Chaos, Solitons & Fractals, 165, 112828. https://doi.org/10.1016/j.chaos.2022.112828

Horie, S., Sato, Y., Furuyama, J., Tanabiki, M., & Aoki, Y. (2019, November). Shot Detection in Racket Sport Video at the Frame Level Using a Recurrent Neural Network. In *2019 15th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)* (pp. 447-453). https://doi.org/10.1109/SITIS.2019.00078

Huang, Y. C., Liao, I. N., Chen, C. H., İk, T. U., & Peng, W. C. (2019, September). TrackNet: A deep learning network for tracking high-speed and tiny objects in sports applications. In *2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)* (pp. 1-8). IEEE. https://doi.org/10.1109/AVSS.2019.8909871

Javadiha, M., Andujar, C., Lacasa, E., Ric, A., & Susin, A. (2021). Estimating player positions from padel high-angle videos: Accuracy comparison of recent computer vision methods. *Sensors*, *21*(10), 3368. https://doi.org/10.3390/s21103368

Javadiha, M., Andujar, C., & Lacasa, E. (2022). A Query Language for Exploratory Analysis of Video-Based Tracking Data in Padel Matches. *Sensors*, *23*(1), 441. https://doi.org/10.3390/s23010441

Lacasa, E., Salas, C., & Torrents, C. (2021). Pádel: una mirada compleja, dinámica y no lineal en la iniciación deportiva y el entrenamiento. *Retos: nuevas tendencias en educación física, deporte y recreación*, (41), 354-361. https://doi.org/10.47197/retos.v0i41.81320

Martín-Miguel, I., Escudero-Tena, A., Muñoz, D., & Sánchez-Alcaraz, B. J. (2023). Performance analysis in padel: A systematic review. Journal of Human Kinetics. https://doi.org/10.5114/jhk/168640

Ramón-Llín, J., Guzmán, J., Martínez-Gallego, R., Vučković, G., Muñoz, D., & Sánchez-Alcaraz, B. J. (2021). Comparison of service tactic formation on players' movements and point outcome between national and beginner level padel. PLOS ONE, 16(10), e0250225. https://doi.org/10.1371/journal.pone.0250225

Sánchez-Alcaraz, B.J., Cánovas, J., Sánchez-Pay, A., & Muñoz, D. (2022). Research in paddle tennis. Systematic review. Padel Scientific Journal, 1(1), 71–105. https://doi.org/10.17398/2952-2218.1.7.

Sánchez-Alcaraz, B., Jiménez, V., Muñoz, D., & Ramón-Llin, J. (2021). Diferencias en los parámetros de carga externa entre el pádel masculino y femenino profesional. Journal of Sport & Health Research, 13(3).

Su, Y., & Liu, Z. (2018). Position detection for badminton tactical analysis based on multi-person pose estimation. In 2018 14th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD) (pp. 379-383). IEEE. https://doi.org/10.1109/FSKD.2018.8686917

Voeikov, R., Falaleev, N., & Baikulov, R. (2020). TTNet: Real-time temporal and spatial video analysis of table tennis. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (pp. 884-885). https://doi.org/10.1109/CVPRW50498.2020.00450

Wu, J., Liu, D., Guo, Z., Xu, Q., & Wu, Y. (2021). TacticFlow: Visual analytics of ever-changing tactics in racket sports. *IEEE Transactions on Visualization and Computer Graphics*, *28*(1), 835-845. https://doi.org/10.1109/TVCG.2021.3114832

Yan, F., Kostin, A., Christmas, W., & Kittler, J. (2006, June). A novel data association algorithm for object tracking in clutter with application to tennis video analysis. In 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06) (Vol. 1, pp. 634-641). IEEE. https://doi.org/10.1109/CVPR.2006.36