

Table tennis ball spin estimation with an event camera

Thomas Gossard* Julian Krismer* Andreas Ziegler Jonas Tebbe Andreas Zell
Cognitive Systems Group, University of Tübingen, Germany

thomas.gossard@uni-tuebingen.de

Abstract

Spin plays a pivotal role in ball-based sports. Estimating spin becomes a key skill due to its impact on the ball's trajectory and bouncing behavior. Spin cannot be observed directly, making it inherently challenging to estimate. In table tennis, the combination of high velocity and spin renders traditional low frame rate cameras inadequate for quickly and accurately observing the ball's logo to estimate the spin due to the motion blur. Event cameras do not suffer as much from motion blur, thanks to their high temporal resolution. Moreover, the sparse nature of the event stream solves communication bandwidth limitations many frame cameras face. To the best of our knowledge, we present the first method for table tennis spin estimation using an event camera. We use ordinal time surfaces to track the ball and then isolate the events generated by the logo on the ball. Optical flow is then estimated from the extracted events to infer the ball's spin. We achieved a spin magnitude mean error of 10.7 ± 17.3 rps and a spin axis mean error of $32.9 \pm 38.2^\circ$ in real time for a flying ball.

1. Introduction

Spin is crucial in ball sports, enhancing ball control and adding an element of unpredictability for the opponent. Whether it is topspin in tennis, backspin in table tennis, or spin variations in cricket, it adds complexity to the game. In soccer and baseball, players use spin to curve shots, making it challenging for opponents to predict the ball's trajectory. This makes spin estimation decisive for playing such sports. Thus, it is also a key component for developing autonomous agents and sports analysis systems. However, in most cases, the spin is too fast for low frame rate cameras, which results in severe motion blur. Event-based cameras do not suffer as much from motion blur as standard frame-based cameras. Indeed, event cameras capture logarithmic changes in a scene's brightness with high temporal resolution in the order of μs . Unlike traditional frame-based cam-

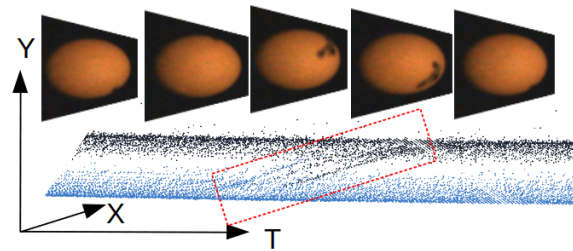


Figure 1. Comparison between the event stream of an event-based camera and captured frames from a frame-based camera. Black and blue dots are, respectively, ON- and OFF-events. The dotted red box highlights the events generated by the spinning logo.

eras, events are only emitted when there is a visual change, offering low latency and reduced bandwidth requirements. This leads to much less information to be transmitted, given a scene with a static background. As a consequence, we do not need to limit the Region of Interest (ROI) of the camera to avoid bandwidth issues. This was indeed a problem in [39], where the full resolution of the camera (1920x1200 pixels) could not be used while maintaining a high enough frame rate (380 fps). Instead, a ROI of 1920x400 pixels was chosen. The ball is thus observed for a shorter time, which decreases the likelihood of observing the ball's logo for low spin values. The event camera can use its full resolution because it does not need to transmit redundant information. The high temporal resolution of event cameras has an additional advantage. Estimating spin using a frame-based camera is constrained by the Nyquist-Shannon sampling theorem, which stipulates that the sampling rate must be at least twice the signal's bandwidth. In other words, we need a camera with a certain frame rate to be able to measure high spin: $\text{fps} > |\omega|/2$, where ω is the ball spin in rps. In Fig. 1, we can see that we only capture two frames where the logo is visible. This makes it extremely difficult to estimate spin accurately. On the other hand, the event camera could continuously capture the movement of the logo. Occasionally, the ball rotates in a manner that causes the logo to consistently appear on the opposite side when captured by a frame-based camera. This is no issue for event-based cam-

*equal contribution

This research was funded by Sony AI.

eras, as events are generated asynchronously. This overcomes the previously mentioned issue as long as the logo moves in front of the camera. For all the previously mentioned reasons, event cameras are better suited for table tennis ball spin estimation.

Contributions of this work are as follows:

- Tracking the ball in real-time and extracting events generated by the logo.
- A real-time and accurate method for table tennis ball spin estimation.
- An evaluation using a ball spinner to show the potential of the approach and a deployment with a ball thrower to show the approach in a real setup.

2. Related work

2.1. Spin estimation

There are multiple approaches for spin estimation in table tennis. Human players mostly estimate the spin applied to the ball from the motion of the opponent's racket. This can be reproduced by using human pose estimation [26, 31], racket tracking [18], or an inertial measurement unit (IMU) installed in the racket [5, 6]. However, the estimated spin is only an approximation, as there is no information on the racket's bounce dynamics, which are greatly influenced by its rubber type. For a more quantitative spin estimate, there are two main approaches: observing the trajectory or directly observing the ball. Trajectory-based spin estimation [11, 34, 39, 42] makes use of the Magnus effect, which bends the ball's trajectory depending on the magnitude and type of spin applied. Nevertheless, these approaches rely heavily on precise estimations of ball positions. The Magnus effect correlates directly with the level of spin imparted on the ball, complicating the differentiation between the spin-induced curvature of the trajectory and measurement inaccuracies of the ball positions. While [2] showed that the ball's trajectory is insufficient for accurate spin estimation, the estimation of the spin can be further improved by using the bounce behavior of the ball [42].

However, for the most accurate spin estimation, ball observation methods are preferred. Using the blur generated by the logo for spin estimation has been studied in [7] but it does not seem to scale well to higher spin values. Within the ball observation methods, we distinguish two sub-categories: logo-based [22, 39, 43] and pattern-based [16, 23, 35, 37, 38, 41]. The patterns are always visible in contrast to the logo which is barely visible or hidden from a single camera's point of view, making previous logo-based approaches unreliable. For official matches, custom balls are not allowed. This makes logo-based methods necessary for competitive games. Our proposed approach leverages event cameras for ball spin estimation.

All of the previously described methods are not exclu-

sive. They can be combined for better spin estimation. A more recent work [36] combined a logo-based method with a trajectory-based approach. The ambiguity about the spin magnitude stemming from the fact that the Nyquist-Shannon sampling theorem might not be fulfilled was resolved by estimating the ball trajectory for various spin magnitudes and choosing the spin magnitude for which the current ball trajectory fits best to the estimated one.

2.2. Event cameras

Event cameras offer many advantages, such as a high temporal resolution and a high dynamic range. However, they also change the processing paradigm. Most computer vision algorithms developed in the last decades are either ill-adapted to process events or do not make the most out of said advantages. Frames can be reconstructed from events to apply state-of-the-art computer vision algorithms, but this introduces extra latency and does not fully utilize the event cameras' speed.

Tracking fast-moving objects is a perfect use case for event cameras, thanks to their high temporal resolution and less severe motion blur. A lot of research has focused on fusing the data from frame-based and event-based cameras [8–10, 44] to compensate for each other's weaknesses: frames for static or slow moving objects and events for fast-moving objects.

Because an event camera generates only ON/OFF events, new algorithms were developed to distinguish different moving objects in a scene. [14, 15] tackle this issue by segmenting events generated by the background and events generated by independently moving objects with a motion-compensated mean time stamp image. Events are then clustered using DBSCAN. Such methods were used for either obstacle avoidance for drones [14] or for capturing a ball with a quadruped robot [15]. Next to the rather specific motion-compensated mean time stamp image, other common event representations are accumulated event frames and time-surfaces [17]. However, the same object will not be viewed the same way depending on its velocity. This was solved with an ordinal time surface such as TOS [21] and EROS [19]. These ordinal time surfaces capture the edge of objects without any trail of events, independent of the object's velocity. A similar representation could be achieved with accumulated event frames, but only if the accumulation time was tuned to fit the object's velocity.

Feature tracking can also benefit from using event cameras, as shown in [3, 12]. However, because of the logo's small size and rapidly changing shape due to the curve of the spinning ball, such methods are difficult to apply.

Optical flow is an alternative for estimating the logo's speed. There are multiple ways to infer the optical flow. There are Lucas-Kanade [27], plane-fitting [4], and time surface matching [28], just to name a few. However, the

current state-of-the-art is learning-based [20]. The most common architectures are U-Net [30], FireNet [32] and RAFT [40]. The optical flow can be learned in either a supervised [20], semi-supervised [24], or unsupervised [25] fashion. Although more accurate, learning-based optical flow methods require more computational power, which makes achieving real-time performance difficult. For this reason, we favor a model-based approach.

3. Method

Our table tennis ball spin estimation pipeline is divided into three phases. Initially, a ball tracker, outlined in Sec. 3.1, estimates the ball’s position, velocity, and radius. Next, this information is used to extract events generated by the logo, as described in Sec. 3.2. Finally, the ball’s spin is estimated from the extracted events, detailed in Sec. 3.3.

3.1. Ball tracker

We set up a static event camera, observing a primarily static background (a table tennis table). This configuration is beneficial since the flying ball triggers most of the events.

We use the Exponential Reduced Ordinal Surface (EROS) [19] event representation for processing the events. EROS works by having new events decay the values of surrounding pixels by a factor, thus enabling objects to have sharp edges, whatever their velocity. This representation enables continuous and asynchronous updates from the event stream, maximizing the potential of the event camera. The update of EROS depends on a parameter named k_{eros} , which represents the update zone around the event. k_{eros} was hand-tuned for the best circle detection and set to 10. A comparison between the accumulated event frames and EROS can be seen in Fig. 2. One issue with the EROS time

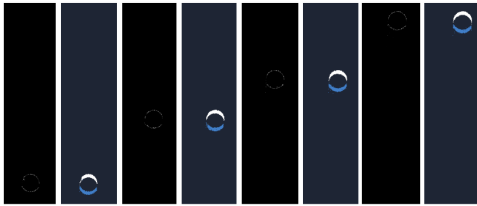


Figure 2. Pairs of representations for the same events. (left) EROS [19] event representations with $k_{eros} = 10$. (right) Accumulated event frames with an accumulation time of $t_{acc} = 2ms$.

surface is the accumulation of noise. Because the noise is spatially and temporally sparse, it stays in the representation. We thus “clean” the EROS time surface by removing isolated pixels. We do so with a hit-or-miss transform to detect isolated non-zero pixels. Detected isolated pixels are then set to zero. This hit-or-miss transform is performed on

every update of the EROS time surface. For more details, we refer to Sec. 9 in the supplementary material.

We do not use convolutions like in [19] because we also need the radius of the ball for extracting the events generated by the logo as explained in Sec. 3.3. Instead, we apply a Hough circle transform to detect the ball and extract the peak value when the accumulation exceeds a certain threshold.

We added a Kalman filter to improve the ball tracker’s performance. This allows for smoother tracking as well as ball velocity estimation. We also estimate the ball’s radius to reduce the search space of the Hough circle transform for faster processing and for converting distances in pixels to meters. The state of the Kalman filter is defined with the state $\mathbf{x} = [x_b, y_b, \dot{x}_b, \dot{y}_b, r]$ where x_b and y_b are the coordinates of the center of the ball, \dot{x}_b and \dot{y}_b are the ball’s velocity in pixels/s and r is the ball’s radius in pixels. Estimating the ball velocity and radius is relevant for extracting the events generated by the logo as explained in Sec. 3.2. The state transition of the Kalman filter is defined as

$$F = \begin{bmatrix} 1 & 0 & dt & 0 & 0 \\ 0 & 1 & 0 & dt & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad (1)$$

where dt is the time difference between two Kalman filter updates. The observation vector of the Kalman filter is $\mathbf{z} = [x_{b,hough}, y_{b,hough}, r_{hough}]$ where $x_{b,hough}$ and $y_{b,hough}$ are the coordinates of the ball center returned by the Hough circle transform and r_{hough} is the corresponding radius. The observation model of the Kalman filter is defined as

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \quad (2)$$

Though the ball’s velocity is not directly observed, it can implicitly be estimated from the sequential measurements of ball positions. The initial variance of the velocities must be set high enough to converge fast enough to the correct velocity. In Fig. 3, we show an example of the Kalman filter tracking the ball. We can observe that the radius of the ball is not constant. This is because the ball can get closer to the camera after being shot.

3.2. Extracting logo events

The ball is tracked accurately using the previously described method. With the Kalman filter, we have an estimate of the ball’s velocity and radius. The ball’s velocity and radius can be used to extract the events generated “inside” the ball, i.e., events generated by the logo. Events that we assume are generated by the logo have to satisfy

$$(y_b(t_e) - y_e)^2 + (x_b(t_e) - x_e)^2 < (r - pad)^2, \quad (3)$$

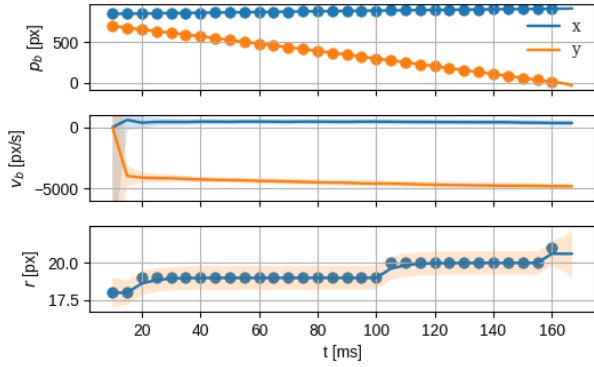


Figure 3. Tracked states of the Kalman filter: Position, $\mathbf{p}_b = [x_b, y_b]$, velocity, $\mathbf{v}_b = [\dot{x}_b, \dot{y}_b]$ and ball radius r . The Kalman filter runs with 200 Hz. The dots indicate measurements from the hough circle transform. The colored band around the estimated values represents the standard deviation.

where t_e is the event's time stamp, x_e the x-position of the event, y_e the y-position of the event, r the radius of the ball and pad is a tolerance to avoid extracting events from the edge of the ball.

For maximal accuracy, the ball's position is continuously updated to the time stamp of the selected event, according to

$$\begin{aligned} \mathbf{p}_b(t_e) &= (x_b(t_e), y_b(t_e)) \\ &= \mathbf{v}_b(t_i) \cdot (t_e - t_i) + \mathbf{p}_b(t_i), \end{aligned} \quad (4)$$

where t_i is the previous time the Kalman filter's state was updated. With this, we can get the events generated by the logo of a flying ball.

3.3. Spin Estimation

Frame-based table tennis spin estimation methods capture successive logo orientations and then regress the spin. Though this method could be translated to event-based cameras, it would not fully exploit their advantages.

Generally, optical flow is a good tool for estimating velocities or, in this case, spin. As mentioned in Sec. 2, there are numerous optical flow methods for event cameras.

We first explain how we obtain the ball's spin ω from the optical flow v_f generated by the logo. From the flow, we can calculate the velocity \mathbf{v} of the point on the ball's surface where the flow is generated

$$\mathbf{v} = \begin{bmatrix} v_{f,x} \\ v_{f,y} \\ \frac{-v_{f,x}e_{r,x} - v_{f,y}e_{r,y}}{e_{r,z}} \end{bmatrix}. \quad (5)$$

We can calculate the z component of \mathbf{v} by knowing that the velocity is tangent to the ball's surface, $\mathbf{e}_r \cdot \mathbf{v} = 0$. This is sketched in Fig. 4. We can then calculate the rotation vector

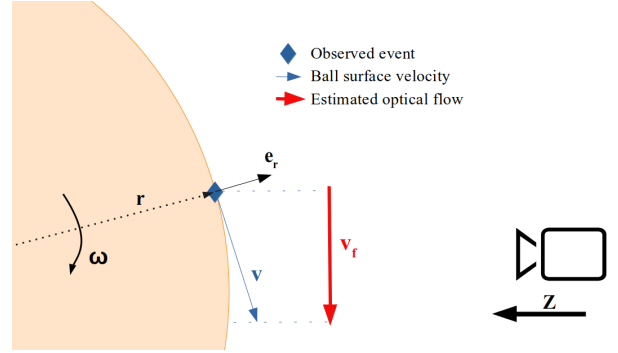


Figure 4. Sketch showing how the spin is calculated from the optical flow.

from the cross-product of the radius and the velocity.

$$\omega = \frac{\mathbf{e}_r \times \mathbf{v}}{\|\mathbf{r}\|}. \quad (6)$$

When working with different optical flow methods, we noticed that they are able to infer flow correctly when the logo is in the center of the ball. But they fail as soon as the logo is on the edge of the ball. This is due to multiple factors. First, table tennis ball logos have multiple edges. As such, events are simultaneously generated at different places of the logo and not in a straight line. This makes identifying successive events generated by the same edge difficult. Secondly, because of the geometry of the ball, the logo's velocity is not constant. There are several parameters to tune the different optical flow algorithms. The main one is the accumulation time, t_{acc} , which aggregates events into an event frame used for optical flow. To obtain the best optical flow results, the t_{acc} should be chosen based on the logo's speed to ensure enough change to observe flow but not so much as to cause motion blur. For a fast, first estimate, we choose t_{acc} as a constant value, based on the median of the expected spin speeds as $t_{acc} = \lceil (\omega/10) \rceil$. From testing, we divided each rotation into 10 time slices because it ensures that even if the true spin speed is higher than we estimated, we still observe the flow multiple times per rotation.

For a better estimate of the spin magnitude and therefore t_{acc} , we also consider a second approach based on the variation of the event rate from the extracted logo events. Indeed, peaks in the event rate indicates when the logo is periodically the most visible. This approach is inspired by [29] and it returns results after a longer observation time but with a higher accuracy. We first apply an exponential moving average (EMA) and subtract the mean event rate. Then, we apply a low pass filter to smooth the curve. After that, we locate all the time stamps at which the sign of this curve transitions from positive to negative values. Since the change in the event rate is periodic and depends on the

ball’s spin, as the logo appears and disappears, it can be used to estimate the magnitude of the ball’s spin. Figure 5 shows how the sign transitions of the event rates EMA can be used to calculate the magnitude of the spin, by taking the time difference between adjacent time stamps, calculating the mean, and converting this time per rotation into the number of rotations per second.

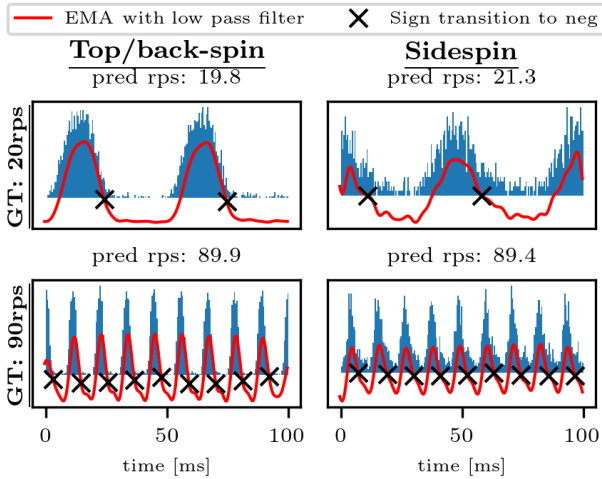


Figure 5. Prediction of the spin magnitude from the event rate. The histograms of the event rate are displayed. We first apply an exponential moving average (EMA) and subtract the mean event rate. Then, we apply a low pass filter to smooth the curve. After that, we locate all the time stamps at which the sign of this curve transitions from positive to negative values.

We can then apply the various optical flow methods for both methods of choosing t_{acc} . The parameters are chosen for each one depending on t_{acc} to return the best results.

For each t_{acc} , several flows and consequently spins are estimated, with the resulting value being the mean of all computed ω . This process can be repeated for each t_{acc} to generate additional spin estimations, which are then averaged collectively.

4. Experiments

We start this section by describing our experimental setup in Sec. 4.1. In Sec. 4.2 we evaluate the ball detector and in Sec. 4.3 we describe our experiment to evaluate the spin estimation of our proposed approach.

4.1. Setup

We use a Prophesee EVK4 event camera with a resolution of 1280x720 pixels for capturing events. The camera is mounted $\sim 2\text{m}$ above the table tennis table, equipped with a 8mm lens. This allows us to cover the whole width of the table, as shown in Fig. 6. The camera’s focus was adjusted using the protocol provided by Prophesee [1]. Event-

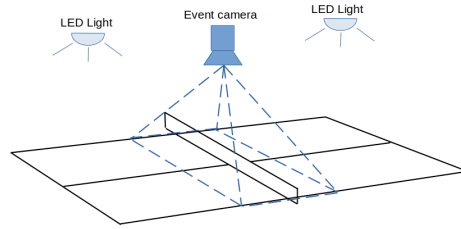


Figure 6. Event camera recording setup

based cameras possess particular settings that differ from standard, frame-based cameras. These were investigated and tuned to generate the most events related to the ball and logo observation while minimizing noise. For more information about the bias settings used with quantitative and qualitative results, please refer to Sec. 6 in the supplementary material. We use LED panels to light the scene uniformly from the ceiling, leading to a luminosity of around 1930 Lux measured on the table. This is to avoid the flickering light from standard lighting powered with 50Hz AC, which triggers events over the entire scene in the event camera, though events generated by the flickering light can also be filtered out with a band-cut filter.

We rely on a custom-made static ball spinner and a ball thrower (Amicus Prime from Butterfly) to generate recordings with known ground truth. The ground truth spin can be calculated from the ball spinner by knowing its relative position to the event camera. We mounted the balls with different logo orientations on the ball spinner to study its impact on the spin estimation, as shown in Fig. 7.



Figure 7. Balls with different logo orientations

For the ball thrower, we rely on SpinDOE [23] to generate ground truth spin, using balls with a custom pattern drawn on them. A frame camera, the Grasshopper3 GS3-U3-23S6C (1900x400), was installed next to the event camera for capturing frames of the balls.

4.2. Ball detection

To evaluate the ball tracker, we recorded multiple balls flying in front of the event camera with speeds ranging from 4 m/s to 12 m/s. The ground truth was generated automatically with blob detection applied on accumulated event frames with a large accumulation time. In Fig. 8, we show

our ball detector’s Mean Absolute Error (MAE) for different ball thrower settings (arbitrary unit). We can see that we have sub-pixel accuracy, which is necessary for extracting the events generated by the ball’s logo.

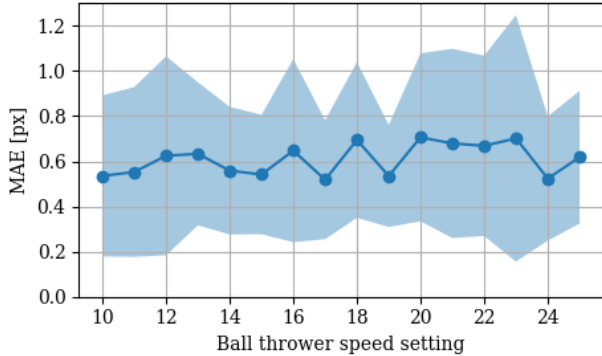


Figure 8. Evaluation of our ball detector for different ball thrower settings (integer values), reporting the MAE in pixels. The colored band around the error represents the standard deviation.

4.3. Spin estimation

We tested our spin estimation method by two different means: with a ball spinner and with a ball thrower. The former has the advantage to give us an extremely accurate and reliable ground truth while the later is closer to real use case. Both benchmarks are still quite similar from the way events generated by the logo are extracted from the flying ball as explained in Sec. 3.2. Indeed, the extracted events’ position is recalculated using the ball position as a reference.

We selected Triplet Matching [33] as the optical flow algorithm. Triplet Matching is the fastest optical flow method we found. It has a sub-millisecond run-time while maintaining competitive accuracy. Moreover, it runs on a CPU, which limits hardware requirements.

We distinguish two types of spin: backspin/topspin and sidespin. Because the ball is observed from the top, backspin, and topspin are almost the same, if only mirrored. With top and backspin, the logo will come into view of the camera and then disappear again. Sidespin balls, on the other hand, will either always have their logo hidden or always visible.

4.3.1 Ball spinner benchmark

We used table tennis balls with different logo orientations to evaluate our proposed spin estimation approach, as shown in Fig. 7. Next to the spin magnitude, our method also estimates the spin axis. The spin magnitude estimation error and the spin axis estimation error with the ball spinner for different ball types are shown in Tab. 1.

Entries with “-” indicate that we did not get any results. For ball 1, the sidespin could not be estimated since the logo and spin axis coincide. The logo of balls 6 and 7 is a rotated version of ball 1; therefore, the same restriction applies.

The errors with the ball spinner for different ball spins are reported in the top part of Fig. 9. As can be seen in the plot, the MAE for the lowest spin magnitude of 10rps is very high. The reason is twofold. First, due to the relatively slow rotation speed, the logo triggers fewer events, which makes it difficult for the optical flow to get accurate flow estimations. Second, with the relatively slow rotation speed, the likelihood of observing the logo over 10ms is relatively low. With 20rps, the MAE is already lower and relatively stable up to 100rps for all four cases except for the sidespin with 10ms observation time. For sidespin with 10ms observation time, the MAE increases linearly with the spin speed. This underlines the difficulty of estimating sidespin.

The MAE of the spin axis estimation errors with the ball spinner for different ball spins are plotted in the bottom graph of Fig. 9. As can be seen in the plot, the spin axis estimation error decreases with higher spin values. This is because higher spin values lead to more events, allowing the optical flow algorithm to achieve higher accuracy.

4.3.2 Ball thrower benchmark

Next to the evaluation with a ball spinner, we also deployed our method in a setup to estimate the spin of flying balls. A ball thrower shot 20 balls. We used velocity settings of 10, 15, 20, and 25 for the ball thrower (respectively approximately 4, 5.5, 7.5 and 9 m/s) and top- and side-spin, with spin strength settings ranging from -5 to 7 . Both our event-based approach and a state-of-the-art frame-based approach [39] were tested on the same observations. The results are listed in Tab. 2. The success rate in the first row is the ratio between the number of successful spin estimations and the total number of ball trajectories. As can be seen, our approach has a success rate slightly above the frame-based approach. Our approach is a bit behind in terms of spin axis and spin magnitude MAE, however, with overlapping error ranges. Given that there is already a considerable body of literature covering frame-based spin estimation methods and our method is the first event-based one, this should show the potential of our approach.

In Fig. 10, we show the spin magnitude MAE on the top and the spin axis MAE on the bottom for a state-of-the-art frame-based approach [39] and our event-based approach. As expected, we see that the frame-based approach has more problems with higher spin settings, while the event-based approach struggles with low spin settings. It is worth noting that though both frame and event cameras have the same FOV, the frame camera boasts a higher resolution ($\sim \times 1.5$). With a resolution similar to the frame camera,

	Ball	1	2	3	4	6	7
10ms	Back/Top-spin	30.0 ± 28.7	27.0 ± 27.5	23.36 ± 26.4	22.4 ± 21.9	32.0 ± 27.3	32.5 ± 28.2
		$18.1 \pm 10.4^\circ$	$25.3 \pm 2.0^\circ$	$27.6 \pm 12.7^\circ$	$25.6 \pm 3.9^\circ$	$15.4 \pm 19.6^\circ$	$9.5 \pm 7.2^\circ$
	Sidespin	-	31.0 ± 20.6	30.5 ± 18.8	32.0 ± 23.4	-	-
		-	$38.17 \pm 14.6^\circ$	$40.3 \pm 14.3^\circ$	$42.8 \pm 7.9^\circ$	-	-
100ms	Back/Top-spin	0.8 ± 0.1	0.5 ± 0.1	2.8 ± 2.7	3.7 ± 3.9	0.6 ± 0.1	0.5 ± 0.1
		$12.5 \pm 4.0^\circ$	$25.1 \pm 1.8^\circ$	$25.7 \pm 1.4^\circ$	$25.5 \pm 1.9^\circ$	$10.7 \pm 5.5^\circ$	$6.2 \pm 3.2^\circ$
	Sidespin	4.8 ± 1.8	8.4 ± 1.5	8.0 ± 3.5	6.5 ± 3.2	-	-
		-	$16.7 \pm 9.4^\circ$	$11.6 \pm 13.5^\circ$	$24.0 \pm 10.5^\circ$	-	-

Table 1. Evaluation of our spin estimation with the ball spinner. Our test was performed with balls with different logo positions relative to the spin axis, as shown in Fig. 7 (white rows: MAE for the spin magnitude in rps, green rows: MAE for the spin axis in degrees). 10ms and 100ms indicate the amount of observation time allowed. Entries with “-” indicate that we did not get any results. For ball 1, the sidespin could not be estimated since the logo only rotates around the spin axis of the ball spinner, and there is no sidespin. The logo of balls 6 and 7 is a rotated version of ball 1; therefore, the same restriction applies.

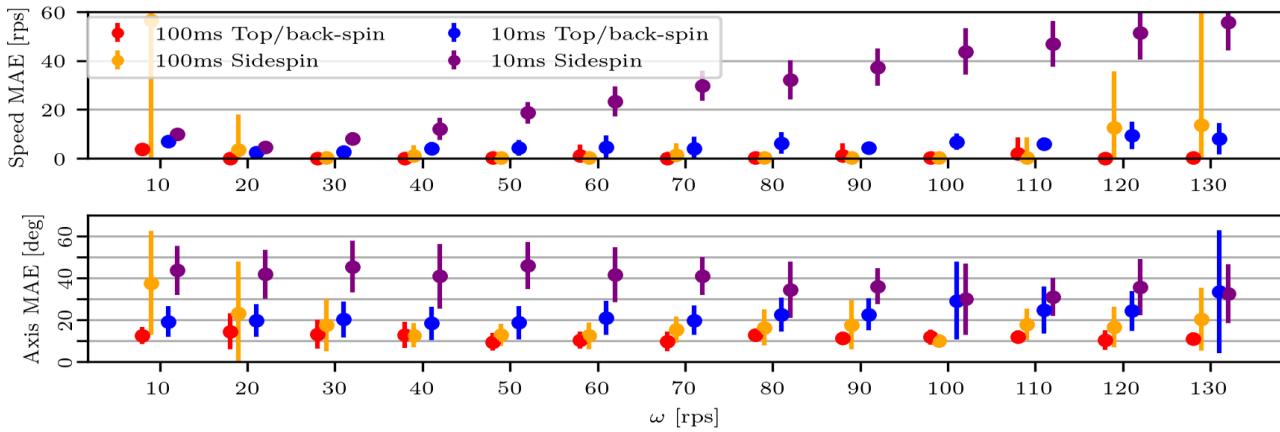


Figure 9. (top) Spin magnitude (speed) error with the ball spinner for different ball spins. (bottom) Spin axis estimation error with the ball spinner for different ball spins. 10ms and 100ms indicate the amount of observation time allowed. Only the events recorded during that time window can be used for the estimation. Backspin/Topspin was evaluated with balls 1, 6, and 7, and sidespin was evaluated with 2, 3, and 4, with an equal number of samples.

	Frame	Event
Success rate	0.79	0.81
Spin axis MAE	$26.6 \pm 33.4^\circ$	$32.9 \pm 38.2^\circ$
Spin magnitude MAE	8.8 ± 15.6 rps	10.7 ± 17.3 rps

Table 2. Benchmark with the flying balls from a ball thrower

the event-based spin estimation would work better. With these complementarity properties, an approach combining an event-based camera and a frame-based camera could be an option.

4.4. Run times

Since our goal is a real time capable spin estimation pipeline, the run-time is another important metric. We used an Intel Core i7-9700 CPU @ 3.00GHz for our experiments,

with all the components written in Python.

From the moment the events are received from the camera, generating the EROS time surface, detecting and tracking the ball, and extracting the logo events takes 12.8 ± 1.2 ms. We run the spin estimation in post-processing and report the ratio between the time it took to process all the events and the duration of the recording. With a ratio of 0.84 ± 0.8 , the spin estimation can process more data than it receives and, therefore, runs in real time.

While it can already be considered capable of being in real time, the setup could be improved. Improving the software design and implementing the software pipeline in C++ instead of Python is expected to boost performance.

4.5. Limitations

While working on the spin estimation for the flying balls from a ball thrower, we noticed fewer events being gen-

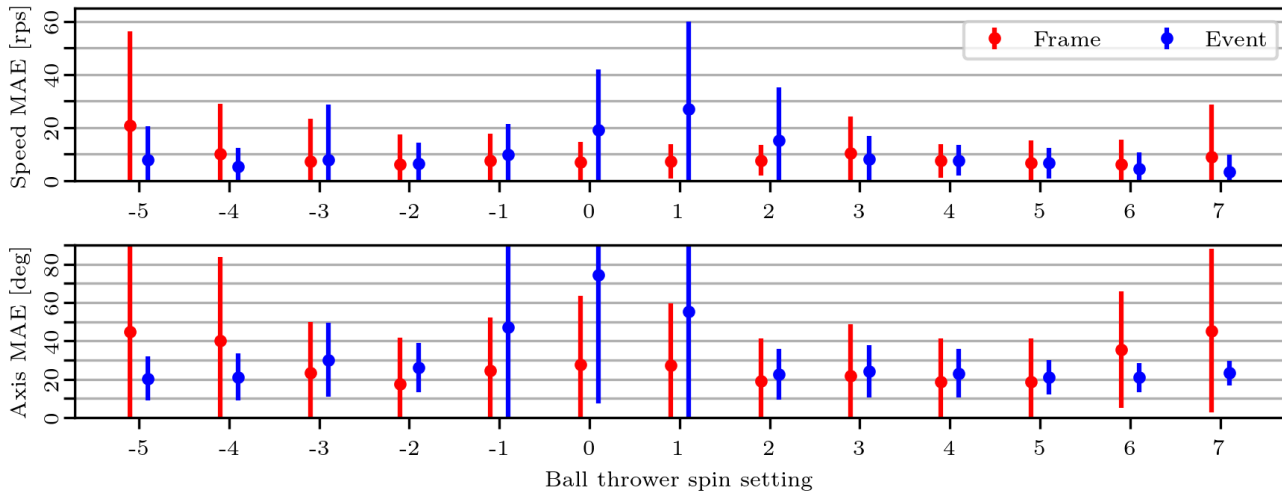


Figure 10. (top) Spin magnitude error with the ball thrower for different ball spins. (bottom) Spin axis estimation error with the ball thrower.

erated compared to the ball spinner setup. This led to an increase in spin estimation failures because there were too few events generated by the logo for the optical flow estimation. Our interpretation of this phenomenon is that pixels respond slower to negative luminosity changes, such as the movement of a logo, after being triggered by a positive change, such as the appearance of the leading edge of a flying ball. This is sometimes referred to as the event camera equivalent of motion blur. Though our method works with the Andro logo displayed in Fig. 11a, it does not generalize to all logos. For example, our spin estimation algorithm does not work with the ball in Fig. 11b since the logo does not generate enough events. This is despite the fact that it works with the ball spinner. We tested different custom logos to determine the causing factor. The logo lines must be thicker than a certain threshold to be detected. We drew lines of different line widths on logoless balls to show this. The ball with the 0.8mm line shown in Fig. 11c did not generate enough events, but the ball with 1.4mm line shown in Fig. 11d worked well with our spin estimation method. As such, only logos with lines of a certain width

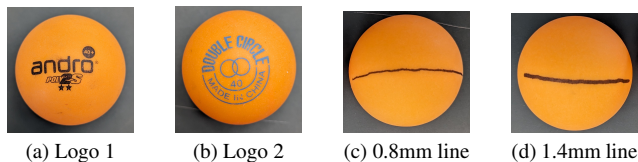


Figure 11. Different ball logos used to estimate the spin of the ball

can be used for our spin estimation with an event camera. Of course, there are different solutions to circumvent this. We could use an event camera with a higher resolution or

a lens with a longer focal length. However, with the latter solution, the camera’s field of view would no longer cover the whole width of table.

We also noticed that the spin estimation has a singularity. Indeed, out of all the balls shown in Fig. 7, there is one logo position, ball 5, for which the spin estimation failed completely, whatever the spin type. Ball 5 does not work for backspin/topspin because the logo is barely visible on the side of the ball. For sidespin, the estimation failed because the rotation axis and the logo coincided. This could potentially be circumvented by using an optical flow algorithm better suited to our specific application.

5. Conclusion

In this work, we presented a spin estimation method for table tennis using optical flow based on the event stream of an event camera. Though this method was implemented and tested for table tennis, it could be generalized to other sports. We were able to compensate for the drawbacks of frame-based cameras with an event-based camera, which allows for more reliable spin estimation.

Although we have shown that our presented approach works for balls on a ball spinner and also for flying balls, depending on the setup, there are still some limitations, notably the slower reponse of OFF events of pixels after being triggered ON.

For future work, we plan to improve our setup and generate a dataset of event recordings of spinning balls to train a neural network specifically for event-based spin estimation. This would allow us to bypass the optical flow estimation. With such a method, we hope to achieve faster and more accurate spin estimation.

References

- [1] Focus adjustment (c++) - metavision sdk docs 4.5.2 documentation. [5](#)
- [2] Jan Achterhold, Philip Tobuschat, Hao Ma, Dieter Buehler, Michael Muehlebach, and Joerg Stueckler. Black-box vs. gray-box: A case study on learning table tennis ball trajectory prediction with spin and impacts. In *Proceedings of the 5th Annual Learning for Dynamics and Control Conference (LADC)*, pages 878–890. PMLR, 2023. [2](#)
- [3] Ignacio Alzugaray and Margarita Chli. HASTE: multi-hypothesis asynchronous speeded-up tracking of events. In *31st British Machine Vision Conference 2020, BMVC 2020*. BMVA Press, 2020. [2](#)
- [4] Myo Tun Aung, Rodney Teo, and Garrick Orchard. Event-based plane-fitting optical flow for dynamic vision sensors in fpga. In *2018 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, 2018. [2](#)
- [5] Peter Blank, Julian Hoßbach, Dominik Schuldhuis, and Bjoern M. Eskofier. Sensor-based stroke detection and stroke type classification in table tennis. In *Proceedings of the 2015 ACM International Symposium on Wearable Computers - ISWC '15*. ACM Press, 2015. [2](#)
- [6] Peter Blank, Benjamin H. Groh, and Bjoern M. Eskofier. Ball speed and spin estimation in table tennis using a racket-mounted inertial sensor. In *Proceedings of the 2017 ACM International Symposium on Wearable Computers*. ACM, 2017. [2](#)
- [7] Giacomo Boracchi, Vincenzo Caglioti, and Alessandro Giusti. *Single-Image 3D Reconstruction of Ball Velocity and Spin From Motion Blur - An Experiment in Motion-from-Blur*. SciTePress - Science and Technology Publications, 2008. [2](#)
- [8] Luis A. Camunas-Mesa, Teresa Serrano-Gotarredona, Sio-Hoi Ieng, Ryad Benosman, and Bernabe Linares-Barranco. Event-driven stereo visual tracking algorithm to solve object occlusion. *IEEE Transactions on Neural Networks and Learning Systems*, 29(9):4223–4237, 2018. [2](#)
- [9] William Chamorro, Juan Andrade-Cetto, and Joan Solà. High speed event camera tracking. In *31st British Machine Vision Conference 2020, BMVC 2020*. BMVA Press, 2020.
- [10] Haosheng Chen, Qiangqiang Wu, Yanjie Liang, Xinbo Gao, and Hanzi Wang. Asynchronous tracking-by-detection on adaptive time surfaces for event-based object tracking. In *Proceedings of the 27th ACM International Conference on Multimedia*. ACM, 2019. [2](#)
- [11] Xiaopeng Chen, Ye Tian, Qiang Huang, Weimin Zhang, and Zhangguo Yu. Dynamic model based ball trajectory prediction for a robot ping-pong player. In *2010 IEEE International Conference on Robotics and Biomimetics*. IEEE, 2010. [2](#)
- [12] Philippe Chiberre, Etienne Perot, Amos Sironi, and Vincent Lepetit. Detecting stable keypoints from events through image gradient prediction. In *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. IEEE, 2021. [2](#)
- [13] Tobi Delbruck, Rui Graca, and Marcin Paluch. Feedback control of event cameras. In *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. IEEE, 2021. [1](#)
- [14] Davide Falanga, Kevin Kleber, and Davide Scaramuzza. Dynamic obstacle avoidance for quadrotors with event cameras. *Science Robotics*, 5(40), 2020. [2](#)
- [15] Benedek Forrai, Takahiro Miki, Daniel Gehrig, Marco Hutter, and Davide Scaramuzza. Event-based agile object catching with a quadrupedal robot. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023. [2](#)
- [16] Sho Furuno, Kazuyuki Kobayashi, Tomoyuki Okubo, and Yosuke Kurihara. A study on spin-rate measurement using a uniquely marked moving ball. In *2009 ICCAS-SICE*, pages 3439–3442. IEEE, 2009. [2](#)
- [17] Guillermo Gallego, Tobi Delbrück, Garrick Orchard, Chiara Bartolozzi, Brian Taba, Andrea Censi, Stefan Leutenegger, Andrew J. Davison, Jörg Conradt, Kostas Daniilidis, and Davide Scaramuzza. Event-based vision: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(1):154–180, 2022. [2](#)
- [18] Yapeng Gao, Jonas Tebbe, and Andreas Zell. *Robust Stroke Recognition via Vision and IMU in Robotic Table Tennis*, page 379–390. Springer International Publishing, 2021. [2](#)
- [19] Luna Gava, Marco Monforte, Massimiliano Iacono, Chiara Bartolozzi, and Arren Glover. Puck: Parallel surface and convolution-kernel tracking for event-based cameras, 2022. [2, 3](#)
- [20] Mathias Gehrig, Manasi Muglikar, and Davide Scaramuzza. Dense continuous-time optical flow from event cameras. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, page 1–12, 2024. [3](#)
- [21] Arren Glover, Aiko Dinale, Leandro De Souza Rosa, Simeon Bamford, and Chiara Bartolozzi. Iuvharris: A practical corner detector for event-cameras. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(12):10087–10098, 2022. [2](#)
- [22] Jared Glover and Leslie Pack Kaelbling. Tracking the spin on a ping pong ball with the quaternion bingham filter. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2014. [2](#)
- [23] Thomas Gossard, Jonas Tebbe, Andreas Ziegler, and Andreas Zell. Spindoe: A ball spin estimation method for table tennis robot. In *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2023. [2, 5](#)
- [24] Woobin Im, Sebin Lee, and Sung-Eui Yoon. *Semi-supervised Learning of Optical Flow by Flow Supervisor*, page 302–318. Springer Nature Switzerland, 2022. [3](#)
- [25] Rico Jonschkowski, Austin Stone, Jonathan T. Barron, Ariel Gordon, Kurt Konolige, and Anelia Angelova. *What Matters in Unsupervised Optical Flow*, page 557–572. Springer International Publishing, 2020. [3](#)
- [26] Kaustubh Milind Kulkarni and Sucheth Shenoy. Table tennis stroke recognition using two-dimensional human pose estimation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pages 4576–4584, 2021. [2](#)
- [27] Bruce D. Lucas and Takeo Kanade. An iterative image registration technique with an application to stereo vision (darpa).

- In *Proceedings of DARPA Image Understanding Workshop (IUW '81)*, pages 121–130, 1981. 2
- [28] Jun Nagata, Yusuke Sekikawa, and Yoshimitsu Aoki. Optical flow estimation by matching time surface with event-based cameras. *Sensors*, 21(4):1150, 2021. 2
- [29] Bernd Pfrommer. Frequency cam: Imaging periodic signals in real-time, 2022. 4
- [30] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. *U-Net: Convolutional Networks for Biomedical Image Segmentation*, page 234–241. Springer International Publishing, 2015. 3
- [31] Soichiro Sato and Masaki Aono. Leveraging human pose estimation model for stroke classification in table tennis. In *Working Notes Proceedings of the MediaEval 2020 Workshop, Online, 14-15 December 2020*. CEUR-WS.org, 2020. 2
- [32] Cedric Scheerlinck, Henri Rebecq, Daniel Gehrig, Nick Barnes, Robert E. Mahony, and Davide Scaramuzza. Fast image reconstruction with an event camera. In *2020 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE, 2020. 3
- [33] Shintaro Shiba, Yoshimitsu Aoki, and Guillermo Gallego. Fast event-based optical flow estimation by triplet matching. *IEEE Signal Processing Letters*, 29:2712–2716, 2022. 6
- [34] Hu Su, Zaojun Fang, De Xu, and Min Tan. Trajectory prediction of spinning ball based on fuzzy filtering and local modeling for robotic ping-pong player. *IEEE Transactions on Instrumentation and Measurement*, 62(11):2890–2900, 2013. 2
- [35] Alexander Szep. Measuring ball spin in monocular video. In *Proc. 16th Comput. Vis. Winter Workshop*, pages 83–89. Citeseer, 2011. 2
- [36] Sho Tamaki, Satoshi Yamagata, and Sachiko Hashizume. Spin measurement system for table tennis balls based on asynchronous non-high-speed cameras. *International Journal of Computer Science in Sport*, 23(1):37–53, 2024. 2
- [37] Toru Tamaki, Takahiko Sugino, and Masanobu Yamamoto. Measuring ball spin by image registration. *Proc. 10th Frontiers of Computer Vision*, pages 269–274, 2004. 2
- [38] Toru Tamaki, Haoming Wang, Bisser Raytchev, Kazufumi Kaneda, and Yukihiko Ushiyama. Estimating the spin of a table tennis ball using inverse compositional image alignment. In *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2012. 2
- [39] Jonas Tebbe, Lukas Klamt, Yapeng Gao, and Andreas Zell. Spin detection in robotic table tennis. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020. 1, 2, 6
- [40] Zachary Teed and Jia Deng. *RAFT: Recurrent All-Pairs Field Transforms for Optical Flow*, page 402–419. Springer International Publishing, 2020. 3
- [41] Christian Theobalt, Irene Albrecht, Jörg Haber, Marcus Magnor, and Hans-Peter Seidel. Pitching a baseball: tracking high-speed motion with multi-exposure images. *ACM Transactions on Graphics*, 23(3):540–547, 2004. 2
- [42] Yuxin Wang, Zhiyong Sun, Yongle Luo, Haibo Zhang, Wen Zhang, Kun Dong, Qiyu He, Qiang Zhang, Erkang Cheng, and Bo Song. A novel trajectory-based ball spin estimation method for table tennis robot. *IEEE Transactions on Industrial Electronics*, page 1–11, 2024. 2
- [43] Yifeng Zhang, Rong Xiong, Yongsheng Zhao, and Jianguo Wang. Real-time spin estimation of ping-pong ball using its natural brand. *IEEE Transactions on Instrumentation and Measurement*, 64(8):2280–2290, 2015. 2
- [44] Yabin Zhu, Xiao Wang, Chenglong Li, Bo Jiang, Lin Zhu, Zhixiang Huang, Yonghong Tian, and Jin Tang. Crsot: Cross-resolution object tracking using unaligned frame and event cameras. *arXiv preprint arXiv: 2401.02826*, 2024. 2