Real-time Fast Moving Object Tracking in Air Hockey

Marcus Philip Rottschäfer[1]

Abstract— **Tracking objects in videos in real-time is a challenging task, especially if the objects that are to be tracked are small and move at fast speed. In air hockey games, tracking the air hockey puck is such a difficult problem. Still, for an air hockey playing robot, estimating the position and trajectory of a puck from video data is the only viable approach. In this paper, we present three real-time approaches to air hockey puck tracking based on computer vision only. We provide a detailed introduction and a critical evaluation of the approaches, aiming at selecting an algorithm most suitable to be implemented in an actual air hockey robot. The approaches are chosen to differ distinctly in the underlying principles and assumptions. This is to give a broad view of various puck tracking principles. For example, two approaches use the pucks location in multiple consecutive frames to derive the pucks trajectory, while another approach solely relies on the motion blur visible in a single image to estimate the trajectory. Next to the motion blur based approach, we present one approach using classical computer vision algorithms and another one based on a deep convolutional neural network object detector. Our results show that out of the three algorithms, there is no such thing as the best approach to air hockey puck tracking caused by a huge variety of gameplay situations in which one approach may perform better than the other two.**

I. INTRODUCTION

Vision is considered one of the fundamental domains of artificial intelligent systems. In the context of a robotic system, it can unlock access to a huge pool of new capabilities and features based solely on the sensory input of a camera. In specific domains, it is currently also the only feasible way for adequately sensing relevant aspects of the environment. For example in detecting and tracking the puck in an air hockey game in order to strategize and command an air hockey robot, vision is the primarily used input modality [6], [10], [12].

In this paper, we present three approaches to air hockey puck tracking based exclusively on vision. The air hockey puck tracking task is about following various motion-related features of the puck like the position and velocity during the airhockey game. Puck tracking is performed solely based on a video stream captured by a camera filming the air hockey playing field. The ultimate goal is to provide useful features for a robot control unit that derives competitive game playing strategies and motor control.

Tracking a puck during an airhockey match is a difficult task. Among various challenges, the small size of the puck and its fast movement, similarity of the puck to the air hockey sticks and playing field markings and real-time processing requirements are the most prominent. In professional gameplay, a puck can move with a speed of up to 15 m/s [10], creating specific hardware requirements for the camera as well as software requirements for processing speed and tracking quality.

Providing a comparison of various object tracking methods for puck tracking, we therefore limit ourselves to the methods allowing real-time detection of fast moving objects like air hockey pucks. The concept of a Fast Moving Object (FMO) originates from sport video analysis and describes an object that moves over a distance greater than its own size within a single frame. These objects create motion blur in the frame and, commonly appearing in a video stream of air hockey gameplay, provide additional difficulty for the tracking system.

We present three approaches to real-time air hockey puck tracking each based on a different principle. The first approach uses classical computer vision algorithms and hand-designed features to localize a puck and estimate its trajectory based on consecutive frames [6]. Contrary to that, the second approach solely relies on motion blur visible in a single frame to estimate position, trajectory and velocity of the puck [10]. Both approaches fundamentally differ in the way they perform the puck tracking and prediction of its future positions. The third approach attempts to completely spare the need for hand-designed features by performing puck detection with a deep convolutional neural network [12].

After presenting all three approaches to puck tracking, we compare them in terms of real-time inference capabilities, detection and tracking performance as well as advantages and disadvantages regarding our specific air hockey system at hand.

II. BACKGROUND

Object tracking is a prominent task in the field of computer vision. It can be formulated as the problem of finding the trajectories of objects moving through a scene in a video sequence [14]. So, given a set of objects detected in a frame, the task is to follow the movements of the given objects in consecutive frames of the video. While just assigning the tracked objects IDs and following those labels would fulfill that definition, additional object-centric information can be collected by the tracker such as object orientation, area or object shape [14].

Enormous progress has been made by the research community in object tracking in the recent decades, promoted by, but not limited to: availability of more compute, better cameras and algorithmic development, large high-quality

*This work was not supported by any organization

[1]Marcus Philip Rottschäfer is with the Faculty of Mathematics, Informatics and Natural Sciences, Department of Informatics, University of Hamburg, 20146 Hamburg, Germany marcus.rottschaefer@studium.uni-hamburg.de

datasets and the rise of neural architectures for image processing (e.g. [4]). Still, object tracking remains a difficult task. As an example, problems that the object tracker might be confronted with, are: [13], [14]

- complex object motion/shapes
- scene illumination changes
- moving camera
- nonrigid objects
- background clutters
- partial and full object occlusions
- real-time processing requirements

Among these numerous reasons for the difficulty, we focus on the last two specifically. That is, for the task of applying an object tracking system to an air hockey robot, we require the system to be capable of real-time processing and to be robust to partly/fully occluded objects like pucks, for example when the arm of one player conceals the air hockey puck in a specific video frame.

Next to the real-time processing capability, we require our object tracker to work with FMOs. The concept of a FMO is commonly used in the context of sport event video analysis; it is defined as any object in a video that, during a single frames exposure time, moves over a distance greater than the objects own size [11]. These objects typically rotate along an arbitrary rotation axis with an arbitrary angular speed. FMOs are often visible as semi-transparent streaks extending along the direction of the objects movement. Examples for FMOs are found in sports, but also occur in other real-life situations like surveillance, falling things, fireworks and hail (see Figure 1 for some examples). Due to the immense speed

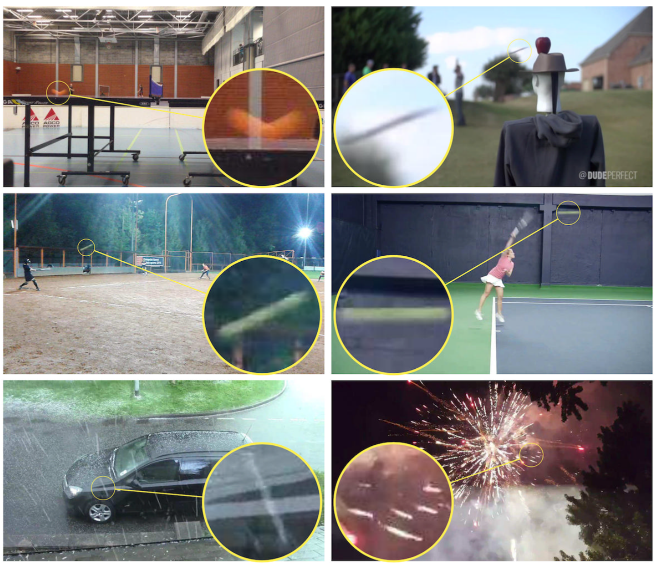

Fig. 1: Examples from video frames where FMOs occur. From left-to-right, top-to-bottom this is table tennis, archery, softball, tennis, hailstorm and fireworks. The FMO is often visible as semi-transparent blur, also called motion blur. Taken from [11].

of a puck (up to 15 m/s [10]), object tracking systems for air hockey must be capable of FMO detection and tracking. Further, velocity and trajectory estimation of the puck are required from the object tracking system.

III. EVALUATION

Localization of FMOs is a tricky task for computer vision. Despite often perfect semantic scene understanding, even humans sometimes struggle in detecting FMOs in images, heavily due to the motion blur caused by the object moving during the frames exposure time [11]. Although often undesirable because of semi-transparence and object-shape obscuring, motion blur is a valuable feature in FMO detection and tracking. This is because the motion blur naturally encodes information about the objects speed and its trajectory [10], [11]. In the following subsections, we evaluate and compare different approaches to air hockey puck localization and tracking, some of which explicitly make use of motion blur.

A. *Classical Computer Vision Approach to Puck Tracking*

Various approaches to puck tracking based on classical computer vision algorithms exist (e.g. [2], [3], [5]), most of which rely on color filtering and hand-picked image features. In this section we want to present one approach to puck tracking based on classical computer vision methods that is comparatively new, fast and especially robust to spots of different brightness.

Li et al. propose a puck tracking system based on hand-coded features and automatic image threshholding [6]. The resulting object detector is able to perform puck detection and trajectory prediction in real-time, on average processing a single frame in just 19.6ms [6].

The detection consists of three steps. Firstly, the RGB-encoded image is converted into the CIELAB color space, the lightness parameter L is set to 50 and the resulting image is converted back into the RGB color space. The authors argue that this step eliminates light interferences and makes the detector more robust to spots of uneven brightness in the playing field.

Secondly, the Otsu algorithm [7] for automatic image thresholding is applied. This gives a binarized image with the puck highlighted as an image segmentation. As the Otsu algorithm assumes input with a single image plane, the authors propose to use the green color plane of the RGB space image as the input. This is a hand-picked choice finetuned on the specifics of the airhockey setup at hand, as their setup is played with a green air hockey puck. So the resulting algorithm is not robust to differently colored pucks, using the algorithm with another puck would require to finetune on the specifics of the new color.

Figure 2 visualizes the result of the preceding two stages. The top-most image depicts the input to the algorithm. The third row shows the result of the CIELAB color space transformation, eliminating spots of uneven brightness in the image. The puck become more blurry and transparent but background illumination can be better accounted for. Finally, the bottom most image depicts the result of the Otsu algorithm based puck detection.

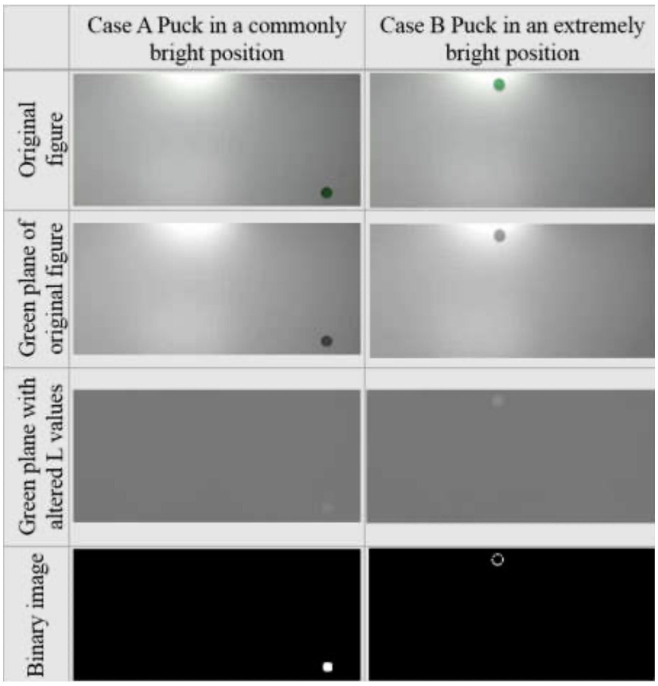

Fig. 2: Processing of two input images from top to bottom, showing the intermediate steps of the pipeline. In the third row, the CIELAB color space transformation has been applied. The bottom most images depict the result of the Otsu algorithm for automatic image binarization. As the two cases indicate, the procedure is relatively robust to spots of different brightness on the playing field. Taken from [6].

The detection pipeline up to now can successfully detect a green air hockey puck on the playing field, however the current approach is still very sensitive to noise. Often the puck is recognized and so are some false positives spread over the playing field. These false positives are small heaps of pixels, often just a single pixel in the binary image. However, they are disturbing the puck localization. Therefore a method is needed to find the puck in the binary image although noisy pixels are present.

The authors propose a method based on the principle of center of mass. The idea is the following:

1) Calculate the center of mass of all detection candidates in the binary image
2) Draw a rectangle of a specific size around that center of mass
3) If more than one object has been found inside the rectangle: Select the largest object within that rectangle and repeat the procedure for the selection.

When only one object has been found inside the rectangle, that object is assumed to be the puck. The puck has then been successfully localized based on the calculated center of mass. Further trajectory prediction is performed as soon as the pucks center of mass has been calculated. The linear trajectory estimation is calculated based on pucks center of mass from the two previous frames. The whole procedure can be performed on a 480p video stream in real time [6]. At no point in time does the algorithm make explicit use of motion blur.

B. Neural Network based Puck Tracking

Neural Network architectures and Deep CNNs specifically have had a huge impact on the field of object detection and tracking in the past years. Tolmacheva et al. build up on that development and propose an approach to air hockey puck tracking solely based on deep learning [12].

For their system, they trained a YOLOv2 object detector [8] on a video dataset of puck positions. In total, the dataset comprises five minutes of airhockey gameplay with the pucks bounding box serving as a ground truth label.

YOLOv2 was chosen due to its real-time inference capabilities and its competitive performance on object detection tasks. Additionally, the authors trained a tiny YOLO model to run on a Rasperry Pi 3 microcomputer.

Figure 3 shows the predictions made by the YOLOv2 network on two example video frames. The network appears to successfully distinguish between the air hockey pucks and the air hockey sticks. This distinction is somehow difficult because both objects have the same color and a similar shape [12]. Although the distinction is good, in the top example image the network wrongly predicts the left pucks position.

As an advantage of the generalization power of the underlying deep neural network, the resulting network robustly deals with varying camera angles and rotations [12].

Different statements where made in terms of the resulting processing speed; at minimum, the authors talk about a performance of 30 FPS, going up to 129 FPS for a different video stream. So getting a reliable performance estimate is difficult; for the specific network variant used in the paper, YOLOv2 416x416 with the Darknet-19 backbone, Redmon et al. report a FPS rate of 67 with a Geforce GTX Titan X [8].

The authors report a "correct detection of the object in the range of 80%" [12]. However, it is not exactly clear how accuracy should be interpreted in the context of an object detection task. It should be noted that possibly the authors talk about the Mean Average Precision (mAP) evaluation metric when they mention accuracy; due to this inconsistency, we will not rely on the evaluation results reported by the authors.

Another YOLOv2 model for puck detection trained by Adleson et al. [1] achieves a mAP score of 0.85 at an IoU level of 0.5. However, their model was trained on a different training set with the intention to compare Deep Neural Network based puck tracking with classical computer vision puck tracking. They report a rapid decrease in detection performance for more accurate IoU levels; for IoU=0.8, their YOLOv2 model achieves a mAP score of just 0.03 in idealistic settings [1]. Assuming that their YOLOv2 model is comparable with the YOLOv2 model trained by Tolmacheva et al., we conclude that accurate neural network based puck detection is difficult or requires better training data than that available for the two models.

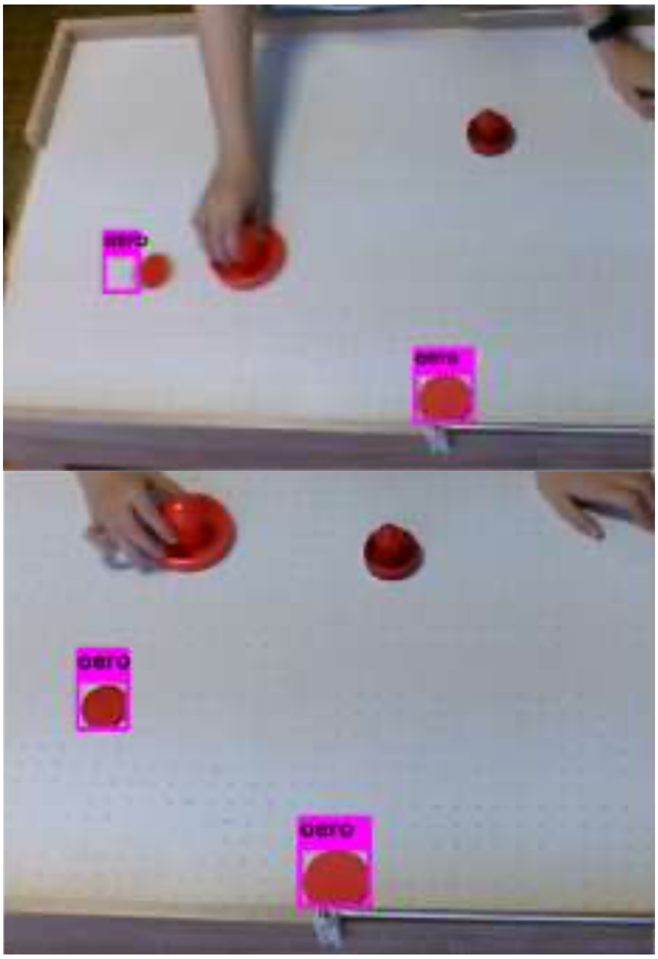

Fig. 3: Puck detection by a YOLOv2 model on two example frames. The network correctly distinguishes between the air hockey sticks and the pucks. However, the detection of the left puck in the upper image is clearly inaccurate. Taken from [12].

C. *Puck Velocity Estimation based on Motion Blur*

Rezvankhah et. al. propose an approach to air hockey puck velocity estimation that relies entirely on the motion blur visible in a video frame [10]. They argue that the estimation with motion blur has some advantages compared to other trajectory and velocity estimation techniques such as those based on consecutive frames.

The first advantage is that a motion blur based approach needs just a single image to perform the estimation. This is because all of the information necessary to derive the velocity are, in theory, available and encoded in the blurry frame. Consecutive frame-based approaches in contrast need at least two frames to perform the estimation, as they try to derive the speed from the difference in position. This is a requirement not always fulfillable with high-speed moving objects. With very fast moving objects, direction changes may occur so quickly that the camera might not be able to capture all changes, thus making precise velocity estimation impossible [10]. As an example in the air hockey context, this can occur when a puck hits one of the walls in between two consecutive frames and hence estimated velocity is incorrect. So, a motion blur approach can process a captured image directly and does not need to wait for a consecutive frame to estimate the velocity.

Additionally, motion blur has an inverse proportional relationship to the frame rate of the camera, in contrast to classical puck tracking approaches: While consecutive frames ideally need a high FPS camera in order to get sharp and unblurry images of the object for accurate position estimation and thus velocity estimation, motion blur approaches benefit from a low FPS camera. This is because the lower the frame rate, the longer the exposure time can be and thus the better and longer the motion blur will be visible in the image.

The approach from Rezvankhah et. al. is based on a hierarchical processing and allows real-time velocity estimation with 30 FPS [10]. It can be subdivided into three steps: puck localization, motion estimation and finally velocity estimation.

1) Locating the puck: For the puck localization and the consecutive steps, first some image preprocessing is applied. Preprocessing performs illumination normalization, resolution reduction, background extraction and grayscale conversion [10].

The resolution reduction is a two-level step, resulting in a lower resolution and a higher resolution image. First, the lower resolution image is use to get a rough estimate of the pucks position, which gives a Region of Interest (RoI) in the higher resolution image. Accurate puck position estimation is derived from the higher resolution image [10]. Using first the rough localization, followed by finer grained estimation based on the proposed region creates a hierarchical procedure and enables the algorithm to do real-time inference.

Grayscale conversion allows the algorithm to work with pucks of any color and gives a small processing speed-up [10].

For the puck localization (motion blur localization), the algorithm uses background subtraction, as well as hand-designed features and heuristics. The result of this step is a grayscale image with just the puck motion blur streak highlighted [10].

2) Estimating the direction of motion: To estimate the direction of the pucks movement, the authors apply a two dimensional Fourier transform to the image window extracted around the motion blur. More specifically, they apply a Fourier transform based filter along a few basic directions and interpolate over all other directions with a mathematical method called steerable filters [9], resulting in the estimation of the motion direction. They do not report on how a Fourier transform allows direction estimation and why they chose this approach. With this method, they achieve an average error of 7.79 degrees to the real direction of motion [10].

3) Estimating the puck velocity: Given the velocity direction, the last task is now to estimate the velocity magnitude. The estimation is based on the following observation: the higher the speed, the higher the number of blurred pixels along the direction of motion [10]. Additionally, the procedure assumes a circular shaped object, which is a valid assumption for an air hockey puck. The authors propose to estimate the distribution of the blurred pixels simply by counting the number of pixels along that direction of motion and to compare that with the distribution of pixels perpendicular to the axis to derive the speed. More specifically, they calculate the speed of the puck from the sharpness difference between the two distributions [10]. Figure 4 shows the two distributions of blurred pixels and the corresponding two orthogonal axes. In the last step, a conversion of the measured speed into a real physical unit (m/s) is performed, based on some constants like table size, resolution, shutter speed, etc. [10].

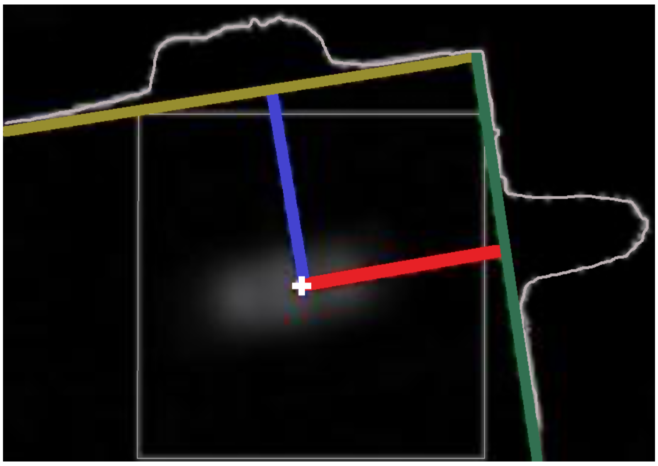

Fig. 4: Estimation of the blurred pixel distributions made by the algorithm. In the middle, the motion blur caused by the puck is visible. The blue line indicates the axis perpendicular to the direction of motion. Inspired by [10].

The algorithm has been implemented in OpenCV and is able to perform real-time processing for 30 frames per second [10].

IV. DISCUSSION

After introducing into three different approaches to air hockey puck tracking presented in the literature, we want to analyse and critically compare the approaches in this section. The algorithms will be discussed in terms of their real-time inference capabilities, detection and tracking quality and applicability to the specific air hockey setup available for our air hockey robot.

A. Real-time Inference Capabilities

All of the approaches presented in the previous section are capable of real-time air hockey puck tracking. Regarding this requirement formulated earlier, we observe that all of the approaches are suitable for our air hockey robot vision system. Still, slight variations exist in terms of the performance. Table I gives an overview of the reported processing speed of approaches III-A to III-C. We observe that with the hardware available to the respective authors, the motion blur based velocity estimation, approach III-C, performs the slowest. The YOLOv2 model is difficult to compare with the other two approaches, as 1) The FPS rate is not reported unambiguously in the paper and 2) It is GPU-based, an acceleration from which the other approaches would, in theory, benefit from heavily. Specific YOLOv2 variants (varying CNN backbone or input size), trading inference speed against detection performance, can achieve up to 91 FPS on state-of-the-art hardware [8]. For the specific network variant used for the neural network puck detection, Redmon et al. report a performance of 67 FPS [8].

TABLE I: Performance of the different approaches in FPS on the authors respective hardware. Note that approach III-B is GPU-accelerated. FPS can not be compared meaningfully.

Approach	FPS	Resolution
III-A [6]	51	480p
III-B [12]	30-129, 67 [8]	416x416
III-C [10]	30	unknown

However, a comparison of the processing speed based on FPS is difficult, as long as no comparable hardware is used by the respective authors. We observe that approach III-B can achieve the highest FPS score, given a powerful GPU. Even with some performance optimization, method III-C is likely to not be applicable to 60 FPS video streams. Though we do not regard this as a problem, since as a motion blur based approach, it is expected to work better with low FPS [10].

What should be taken from this part is that likely all approaches can be used for real-time inferencing, given suitable image resolution and hardware, though just the GPU-based III-B might be able to process video streams with $\geq$ 60 FPS. For this result, we ignore performance optimizations and hardware improvements possibly available to boost the specific algorithms inference speed.

B. Object Tracking Performance

Of all the three approaches, just the authors of the YOLOv2 model provide some objective puck detection evaluation talking about a correct detection in 80% of the cases, as presented in section III-B. As mentioned, the detection performance decreases rapidly for more accurate puck detection levels, following the results of a very similar approach of Adleson et al. [1]. Conveying, for a centimeter or even millimeter-precise puck localization, although fast, the neural network approach is likely not the way to go.

Due to missing evaluation methodology, we can not report on the tracking performance of the other approaches here.

Another point concerning puck tracking performance and transferability to our air hockey setup is derived from some simplifications of the environment made by the authors. The neural network and the classical computer vision approach both do not have any distractions on the tabletop playing field that could interfere with the puck detection. The playing field is just plain white. The air hockey table playing field available to us that serves as our static background on the other hand does have some obstacles, geometric shapes like a mid-table line, goal-lines, side-lines and circles, which are not accounted for in the implementation of the approaches. Performance results collected by the authors may therefore not be directly transferable to our concrete air hockey setup at hand.

Also, the classical computer vision approach and the motion blur approach both do not account for air hockey sticks or hands interfering with the puck detection, e.g. occluding the puck or misleading the detection due to similarity in shape to the air hockey puck. Just the neural network approach is provably robust to possible false positives originating from the similarity of the puck to the air hockey sticks. The authors report that their model successfully learns to distinguish between those two entities [12].

Care must be taken when using an object detector for each frame for the object tracking task. The YOLOv2 model for example will be applied to each video frame individually, disregarding any information about the pucks position that could have been acquired in previous frames. This means that this approach totally ignores the dependency of consecutive video frames with each other. This independence-assumption made by the detector could lead to poor tracking performance. The same drawback applies to the classical computer vision approach.

Just the motion blur based approach uses some information from the short-term image history in a specific video sequence to improve the detection for the current frame. It does so by applying a background subtraction method based on Gaussian Mixture Models [15] incorporating knowledge from a few previous frames. However, the motion blur detection itself does not make use of motion blur features from the previous frames, like the previous motion blur position, to improve the tracking performance.

This is an improvement all of the three approaches could benefit from heavily. It becomes even more important for the neural network and the classical computer vision based approaches, as they perform trajectory estimation based on the pucks position in two consecutive video frames. The trajectory estimation will fail if the pucks position in both frames has not been estimated correctly. Therefore a correct puck detection in two consecutive images is crucial, contrary to the motion blur based approach where just a single image is used for trajectory and velocity estimation.

C. Applicability to our Air Hockey Setup

It is unclear, how the motion blur based approach will perform if no motion blur is visible in the image, that is for example when the puck is moving very slow or a high-FPS rate camera is used. The authors do not report on such cases [10]. We expect the approach to perform miserably for trajectory and velocity estimation when little to no motion blur is available.

On the contrary side, it is unclear how the neural network and especially the classical computer vision approache assuming circular shaped objects will perform when noticeable motion blur is present in the image. Again, the authors do no report on such cases [6], [12]. As both approaches do not make explicit use of motion blur as a feature, we expect both approaches to decline in detection performance when motion blur is present.

For a realistic air hockey game, we found the video frames to have both motion blur and no motion blur, depending on the current pucks speed in the gameplay. A recommendation on which approach to use for our air hockey setup is therefore difficult, as each approach individually offers some advantages in specific gameplay situations. For example, with a slow-moving puck and no motion blur, the classical computer vision approach will likely bring convincing tracking performance, while for blurry images in fast-moving puck situations, the motion blur approach will be better suited for localization and trajectory estimation.

V. CONCLUSION

We presented and evaluated three different approaches to the air hockey puck tracking task. All approaches are distinct in the principles and underlying algorithmic building blocks performing the puck tracking from a given video stream in real-time. While one approach relies on the motion blur visible in a single video frame only, the other two approaches are better suited for video frames where little to no motion blur is present. They are based on classical computer vision algorithms and deep learning, respectively.

In the comparison of those approaches to the puck tracking task, we found that there are distinct advantages and disadvantages inherent in each of the methods, making a general statement about superiority of one approach over the other two difficult.

In terms of processing speed, all approaches are capable of real-time inference at a minimum of 30 FPS, however only the neural network based, GPU-accelerated approach is likely to process 60 FPS video streams or more.

In terms of puck tracking quality, the motion blur based approach is likely to perform best with very fast moving pucks resulting in blurry images, while the other two approaches are better suited for unblurry frames, where the puck moves slowly or is not moving at all. Although the fastest in terms of FPS, the neural network approach was reported to suffer from a low precision when millimeter-precise detection is aimed for. We conclude that for a robust puck tracking system to be used in a real air hockey robot, a combination of those approaches is likely to achieve the best results.

For future work, we will extend our selection of algorithms for puck tracking to include more methodologically-different approaches. For example, in the class of the classical computer vision algorithms, there are many more approaches to be mentioned and to be compared with each other. Additionally, we would like to look into air hockey puck tracking with event cameras, also known as neuromorphic cameras, a new camera technology promising high FPS rates and a better applicability to motion-heavy domains like for example air hockey.

REFERENCES

[1] Christopher Adleson and David C Conner. Comparison of classical and CNN-based detection techniques for state estimation in 2D. *Journal of Computing Sciences in Colleges*, 35(3):122–133, 2019.

[2] Ahmad AlAttar, Louis Rouillard, and Petar Kormushev. Autonomous air-hockey playing cobot using optimal control and vision-based bayesian tracking. In *Annual Conference Towards Autonomous Robotic Systems*, pages 358–369. Springer, 2019.

[3] Martin Ansgar, Hanna Berggren, Pontus Borg, Rasmus Damberg, Mattias Gudjonsson, and Lukas Mared. *Autonomous Air Hockey*. PhD thesis, 2016.

[4] Gioele Ciaparrone, Francisco Luque Sánchez, Siham Tabik, Luigi Troiano, Roberto Tagliaferri, and Francisco Herrera. Deep learning in video multi-object tracking: A survey. *Neurocomputing*, 381:61–88, 2020.

[5] Matthew Lakin. A Robotic Air Hockey Opponent. *CCSC: SC Student Paper E-Journal*, 2, 2009.

[6] You-Cheng Li and Shih-Yen Huang. Fast-moving object tracking in air hockey. In *2017 IEEE International Conference on Mechatronics and Automation (ICMA)*, pages 1693–1697. IEEE, 2017.

[7] Nobuyuki Otsu. A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics*, 9(1):62–66, 1979.

[8] Joseph Redmon and Ali Farhadi. YOLO9000: better, faster, stronger. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7263–7271, 2017.

[9] Ioannis M Rekleitis. Optical flow recognition from the power spectrum of a single blurred image. In *Proceedings of 3rd IEEE International Conference on Image Processing*, volume 3, pages 791–794. IEEE, 1996.

[10] Sh Rezvankhah, Amir Ali Bagherzadeh, Hadi Moradi, and B Nadjar Araabi. A real-time velocity estimation using motion blur in air hockey. In *2012 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pages 1767–1772. IEEE, 2012.

[11] Denys Rozumnyi, Jan Kotera, Filip Sroubek, Lukas Novotny, and Jiri Matas. The world of fast moving objects. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 5203–5211, 2017.

[12] A E Tolmacheva, D A Ogurcov, and M G Dorrer. Puck tracking system for aerohockey game with YOLO2. In *Journal of Physics: Conference Series*, volume 1399, page 33116. IOP Publishing, 2019.

[13] Yi Wu, Jongwoo Lim, and Ming-Hsuan Yang. Online object tracking: A benchmark. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2411–2418, 2013.

[14] Alper Yilmaz, Omar Javed, and Mubarak Shah. Object tracking: A survey. *Acm computing surveys (CSUR)*, 38(4):13–es, 2006.

[15] Zoran Zivkovic and Ferdinand Van Der Heijden. Efficient adaptive density estimation per image pixel for the task of background subtraction. *Pattern recognition letters*, 27(7):773–780, 2006.