Journal of Science & Cycling
Breakthroughs in Cycling & Triathlon Sciences

OPEN ACCESS

Article

Automatic Mapping of Finish Line Videos for the Objective Analysis of Sprint Behavior

Pieter Verstraete [1], Jelle De Bock [1], Arie-Willem de Leeuw [2], Tom De Schepper [2], Steven Latré [2] and Steven Verstockt [1,*]

[1] Ghent University - imec, IDLab; Technologiepark-Zwijnaarde 122, 9052 Gent; pieter.verstraete@ugent.be, jelle.debock@ugent.be, steven.verstockt@ugent.be

[2] University of Antwerp - imec, IDLab; Sint-Pietersvliet 7, 2000 Antwerpen; Arie-Willem.deLeeuw@uantwerpen.be, Tom.DeSchepper@uantwerpen.be,Steven.Latre@uantwerpen.be

* Correspondence: (SV) steven.verstockt@ugent.be.

Received: 13 May 2021; Accepted: 28 May 2021; Published: 30 November 2021

Abstract: This paper proposes a computer vision-based methodology to generate riding line maps from bird's eye view sprint videos. These maps can be used to objectively evaluate dangerous sprint behavior or to perform sprint performance studies. In order to generate the maps, our automatic workflow first extracts the road and riders from the video images using state-of-the-art object detection models. Next, feature points in the remaining part of the images are used to estimate the homography parameters and to stitch the overlapping images into a map of the finish zone. The same homography parameters are also used to reproject the riders onto the sprint map. Based on their positions on the map and the timing info from the video metadata, we get a spatio-temporal description for each riders' sprint. These descriptions are stored in JSON format and can be used for further analysis. As a demonstrator, we present some examples of objective evaluations of dangerous sprint behavior. Those evaluations are based on outlier and overlap detections of the riding lines.

Keywords: sprint analysis; sports data science; race cycling performance; computer vision; similarity metrics

1. Introduction

Most modern-day cycling races end in a sprint. If it's not a bunch sprint, it's a leading group that sprints for the flowers or a classic sprint à deux that decides who gets the kisses [1]. As such, being able to study and improve sprint performance is important for each rider. Dilger and Geyer [2] performed a study investigating the slipstreaming effect of riders during sprints. A rider is slipstreaming if he/she is riding reasonably close to another rider in front resulting in a reduced drag force acting on the rider that is slipstreaming. They theoretically and empirically showed that slipstreaming is a key performance indicator in sprints with sprinters of similar capabilities. The combination of optimally using drag benefits, perfect timing to start its final acceleration and good relative positioning within the bunch are the most important contributors to sprint success. Other contributors are, for instance, the rider's sprinting positioning (standing/sitting), the optimal cadence and power output [1, 7, 8].

Furthermore, and in addition to the actual sprint, analysis of the sprint leadout and preparation in the final kilometers can also indicate, among others, optimal positioning and team strategy in function of the race result. The ultimate challenge still lies in the understanding of why, or why not, a rider did achieve success, and how this can be improved. As detailed sensor information is often not publicly disclosed, and cannot capture all contributors (e.g., positioning, leadout), an analysis of video footage seems worthwhile.

Sprints sometimes are also marred by crashes and dangerous riding behavior, impacting the sprint result of some of the riders. A study of Lybbert et al. [3] investigated the implications of changing the position where the *red flag rule (RFR)* is applied from one kilometer to three kilometers from the finish line. The RFR can be defined as the point from the finish line where incidents or mishaps do not influence a rider's cumulative general classification time. They found that shifting the red flag rule to three kilometers did not significantly reduce the amount of crashes in the final sprint kilometers and even suggested that it might in some circumstances even induce greater risk taking by the sprinters as they *felt protected by the RFR*. This study teaches us that it is still very relevant to investigate the bunch dynamics and possible incidents in the last few hundred meters of the race. This is usually the point where sprinters reach their top speeds. Having tools to objectively study these incidents and visualize them in an easily consumable way can facilitate the work of race jury and improve safety in future races (e.g., by making riders more aware about the impacts of their riding behavior). Furthermore, this technology can help the UCI to become more consistent in its judging of violations of rule 2.3.036. This rule states that "*riders shall be strictly forbidden to deviate from the lane they selected when launching into the sprint and, in so doing, endangering others*". Today, it is still the race jury that will - rather subjectively - decide if a rider has done something wrong. A tool that can assist them by flagging outliers (~ abnormal sprint behavior) and showing them similar historical sprints with its decisions that were taken then, would definitely improve the interpretation of the rule. As such, the tool can also be considered as a kind of second opinion for the race jury.

The main aim of our work is to demonstrate that mapping bird's eye view video images onto a sprint map allows us to generate spatio-temporal data of rider positions that can then be used in performance, storytelling and safety studies. Similar computer vision-based mapping methodologies have been studied in literature [4-6], however, we are the first applying them on bird's eye view sprint videos.

The remainder of this paper is organized as follows. Section 2 introduces our dataset and the general architecture of our mapping tool. Furthermore, it discusses each of the mapping tool's building blocks in more detail: road/rider segmentation, feature extraction, mapping and riding line generation. Subsequently, Section 3 shows some mapping and riding line results and demonstrates how they can be used to objectively evaluate dangerous sprint behavior. Those evaluations are based on dynamic time warping based similarity detection of the riding lines. Next, Section 4 lists our major findings, and the practical application of our tool is discussed in Section 5. Finally, Section 6 lists the conclusions and points out directions for future work.

2. Materials and Methods

2.1 Bird's eye view sprint video dataset

In order to generate accurate sprint maps with rider positions our method needs bird's eys view videos of finish lines - this is the only prerequisite of the proposed approach. A challenging World Tour races dataset of approximately 100 bird's eye view finishing line videos of the past five years was created to test our approach. Since we aim for a generic solution that is widely applicable and can be used with no/minor modifications, it is important that the dataset is representative and covers the majority of finish line types. As can be seen in Figure 1, our dataset contains sprints with different characteristics, e.g., shot in rural/urban environments, with different zoom levels, varying weather/lighting conditions and with/without spectators. An evaluation on this dataset will provide us a good estimate of how it will perform on the majority of recent sprints.

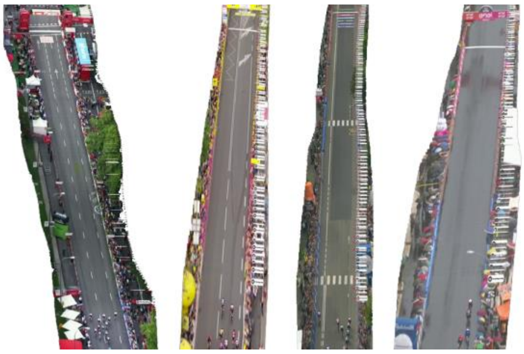

Figure 1. Sprint maps (~ stitched bird's eye view video images) with different characteristics/context - selected from our dataset of 100 recent sprints.

2.2 Sprint map/data generation pipeline

Each bird's eye view finish line sprint video will be analyzed using the pipeline proposed in Figure 2. First, we extract the road and riders from each video image using a Detectron2 based model[1] that is trained to detect these types of objects. Static elements, such as broadcaster logos and timing info, are also masked (and not further taken into account) using a background subtraction method. Once all these objects have been removed, the remaining part of the image is used to find feature correspondences between the image and the sprint map. Different feature point detectors and descriptors have been tested/evaluated. Out of these tests, the Scale Invariant Feature Transform (SIFT) [10] came out as the one with the best performance and accuracy. Based on the feature point matches, we can estimate the geometric transformation parameters to stitch the image to the sprint map. Using the same transformation parameters, rider positions can be transformed to sprint map coordinates. Those coordinates - in combination with the timing info - can be used to study the sprint behavior of each rider. This spatio-temporal rider data is stored in JSON format and can easily be used as input in different types of studies/analyses.

Important to mention is that in order to have an indication of the speed of the riders, the pixel per meter ratio of the sprint map needs to be known. Based on the finish line detection - and more specifically its width (which is fixed in World Tour races as a 4cm black line enclosed within two white bands of 34cm each, as defined by UCI [10]) - we can calculate this ratio. The finish line detection will also allow us to link the rider detections to the stage results and identify which rider corresponds to which detected riding line.

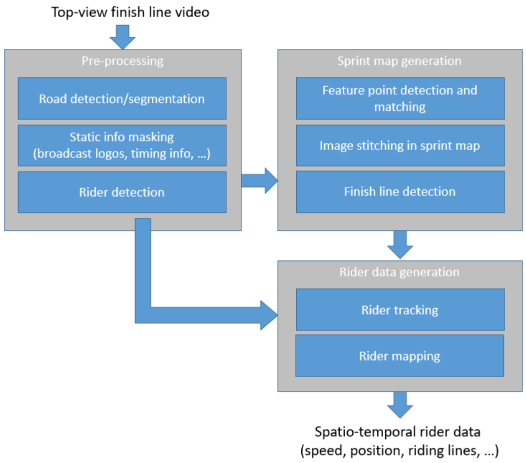

Figure 2. General architecture of the proposed sprint map/data generation pipeline

2.3. Road and rider segmentation

Both the road and rider segmentation make use of Detectron2 (developed by Facebook AI Research and implemented in PyTorch) and have been trained on a dataset of annotated images with road surface and rider segmentations. The image labeling/annotation was done using the VGG

[1] https://github.com/facebookresearch/detectron2

image annotator[2]. In order to train the models, the dataset was also split into a training and test set. All details of the dataset can be found in Table 1.

Table 1. Characteristics of our road and rider segmentation dataset.

Type	#training images	#test images	#training labels	#test labels
Rider	432	95	2139	357
Road surface	432	95	432	95

Both the road and rider instance segmentation are built using the pre-trained mask_rcnn_R_50_FPN_3x model [12]. This model was chosen because it trains fast and has the best speed/accuracy ratio. Learning rate was set to 0.0025 for both classes and the amount of iterations was set to 10000 and 15000 for rider and road objects respectively. The segmentation precision for both classes is above 0.9. Figure 3 and Figure 4 show segmentation results for road and rider segmentation respectively.

Figure 2 3. Road segmentation result

For each detected rider, the bounding box of the segmented region is used to estimate the position of the rider. Currently, the center of the bounding box is used as the rider position - future work will focus on optimizing this, (e.g., by using the center of the rider's helmet as detected rider location). Riders are also tracked across consecutive video frames. This is done by the

Figure 4. Rider segmentation results

SORTalgorithm [11] which uses the position of a rider in the past frames to estimate the new position in the current frame. Based on the estimated and detected positions, each rider is associated with its correct identifier.

2.4. Sprint map generation

The sprint maps are generated iteratively by i) finding SIFT feature correspondences between each new image and the already existing sprint map, ii) estimating the geometric transformation parameters to stitch the image in the sprint map using random sample consensus (RANSAC), and iii) stitching the images in the sprint map. An example is given in Figure 5. Using the same transformation

parameters, rider positions are also transformed to sprint map coordinates.

2.5. Rider data generation

Based on each rider's projected positions on the sprint map, we can visualize the rider's trajectory on the map and generate statistics/insights about it.

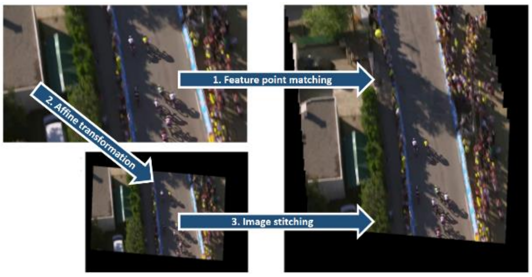

Figure 5. Sprint map generation.

2 https://www.robots.ox.ac.uk/~vgg/software/via/

3. Results

Our main objective was to generate a sprint map on which we can study the spatio-temporal evolution of riders in a sprint. As an example, Figure 6 shows sprint lines for the 2019 Vuelta stage 4 to El Puig.

Figure 6. Sprint map for 2019 Vuelta stage 4 to El Puig

Subjective evaluation of the sprint maps shows that our method definitely has potential to study sprint behavior. However, a decent objective evaluation is needed to correctly measure its correctness/accuracy. We are currently exploring several strategies, such as analyzing the smoothness of the generated riding lines and speed curves, comparisons with Strava sprint results and matching with satellite images. Figure 7 shows the idea behind the latter approach and also highlights some issues that need further research. The skew error and aspect ratio differences between the sprint map and the satellite image, for example, is most probably due to the camera orientation which is not perfectly vertical (i.e., slightly oblique). Post-processing the sprint map using a feature point matching- based transformation with the corresponding satellite image could be a solution to fix these issues.

Figure 7. Comparison of sprint map and satellite image for 2021 Oxiclean Classic Brugge - De Panne.

The speed curves, shown in Figure 8, show the pixel distance of riders between consecutive images for the Giro 2020 stage 3 that finished in Orbetello. As can be noticed, the pixel distance is decreasing when riders get closer to the finish. It is possible that there was indeed a speed drop in the last meters (should be checked in Strava results), but a stitching error (e.g., due to a non-vertical camera angle as discussed before) can also cause such effects. Further research is needed to check what is really going on and will be part of our future work. The quadratic curve fit error for these speed curves is 2.01-pixel distance. This might be further improved by more accurate rider segmentation and localization. The impact of this error on the real speed/location calculations is dependent on the actual real-world distance of a pixel (and the height from where the images were shot).

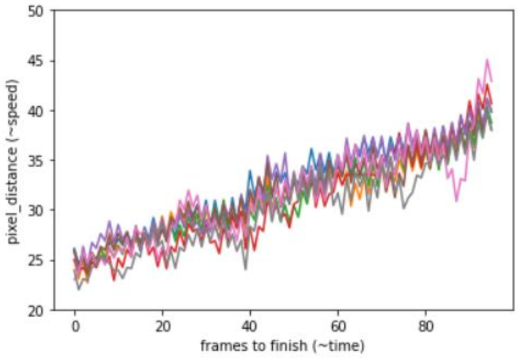

Figure 8. Speed curves for Giro 2020 stage 3 that finished in Orbetello.

4. Discussion

The proposed methodology is a first step towards objective analysis of sprint behavior. Preliminary results show that it performs best on a continuous stream of bird's eye view video images shot at constant speed, fixed angle and with no zooming effects. These instructions can, for example, be given to the broadcast helicopter/VAR to include it into their race coverage/logging protocols. Other possible improvements, such as optimizing the rider's segmentation/localization and improving our evaluation mechanisms, have been discussed in the previous section and will also be part of future work.

5. Practical Applications.

5.1. Objective evaluation of sprint lane deviation

If a detected riding line shows deviating behavior and blocks/impacts another riders' line we can automatically flag it as dangerous riding behavior. This can serve as a trigger for the race jury to further investigate this rider's sprint. As an example, we demonstrate our method on video footage of the Giro 2020 stage 3 that finished in Orbetello. In this stage, the Italian rider Elia Viviani was disqualified from the stage as a result of dangerous sprint behavior. The sprint map for this stage is shown in Figure 9.

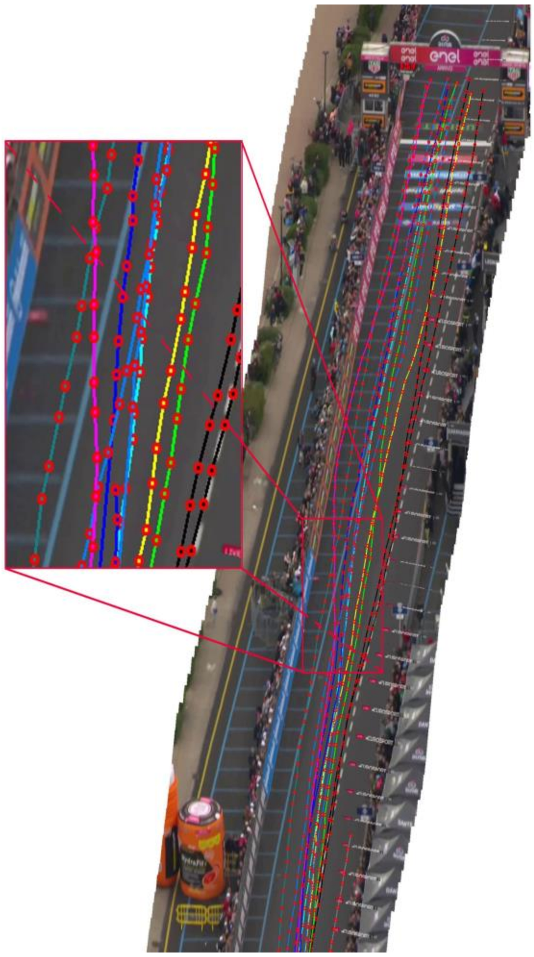

Figure 9. Giro d'Italia - Stage 3 (Orbetello) - Viviani's dangerous riding behavior impacts the riding lines, as can be seen in the abrupt change of them.

Based on the detected rider positions on the sprint map, our algorithm flags riders that change direction and in doing so block the predicted future position of those that are just behind them. The first step of this algorithm consists of predicting the missing rider positions, as not all riders are always present/detected in each frame. We interpolate between the known positions of the rider taking into account the timestamps of those positions. Subsequently, we analyze the riding line angles over time. As can be seen in Figure 10, the riding angles in the Giro 2020 stage 3 indicate an abrupt change for some riders at 50 frames from the finish. Further inspection of those riders' positions and distances to other riders, which is the next step, can then highlight if someone is doing an unallowed manoeuvre. For the Giro 2020 stage 3, we found that one rider (= Viviani) was blocking another rider - his distance to the predicted position of a rider behind him was close to zero, after deviating from his line (as can be seen in Figure 11).

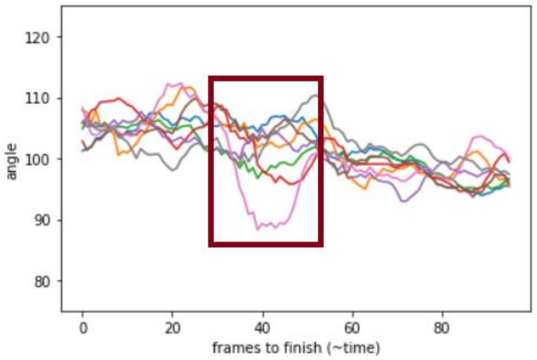

Figure 10. Riding angles of Giro 2020 stage 3. An abrupt change can be detected at 50 frames (~measuring points) from the finish.

Figure 11. Giro d'Italia - Stage 3 (Orbetello) - Viviani's dangerous riding behavior impacts several riders that are just behind him.

5.2. Sprint performance

The second useful application consists of a sprint performance examination. To assess this performance, we transform the data into numerous variables that we believe have an impact on the result of a rider in a sprint. First, we assign a race position to the riders in all frames and by using the position of the finish line, we also determine the final result of each rider. Second, we consider the velocities and accelerations of the sprinters throughout the sprint. Rather than focussing on the absolute values, we will consider relative values of these variables to allow for a comparison of sprints on different terrains. Third, we also consider position variables, such as the position on the road with respect to barriers or position line characteristics.

In future work, we will apply different Machine Learning techniques to explore whether there are patterns in the bunch sprints. Here, the main goal is to find the similarities between the winners of the bunch sprints. In particular, we will investigate which of the aforementioned variables have the greatest impact on the performance.

5.3. Other sports

Line choice and analysis is also useful in other sports/disciplines. In road cycling the sprints are the most suited for the line analysis, but in other disciplines such as cyclocross or cross-country mountain biking other parts from the race can be analyzed as well. In those disciplines, the start is often also important to get a good positioning before the first course feature. The first seconds of these races are usually characterized by lots of position changes and different riding lines spread across the width of the road. Also riding lines in technical features such as sandpits, descents, barriers and technical corners are suited for further analysis. Figure 12 shows an example of such a video tracking analysis of a pro race of the 2021 Cyclocross season. As can be observed, the green riding line is exiting at the other side of the track (compared to the blue and orange riding line). With a transformation matrix to map video coordinates on *real life* coordinates, an indicative speed for each of the riding lines could be provided. The example shown in Figure 12 is the result from a single camera source analysis, but by combining and stitching video frames from multiple camera sources positioned across the segments more detailed course analyses could also be performed. Similar importance of the start can also be found in BMX and Motocross where the riders typically start from behind starting gates. Often those first tens of meters of the track, typically downhill off a ramp or slope, are crucial for the remainder of a race. Hence, rider line analysis might also be relevant here, especially in combination with the typical high speeds that are achieved. Comparable conditions can also be found at the start of Formula One races. Those first rounds tend to be quite hectic with a lot of overtaking maneuvers that quite often result in crashes in the early stages of the race. Finally, line analysis can be applied in several skiing or snowboard disciplines, where optimal cornering is very important.

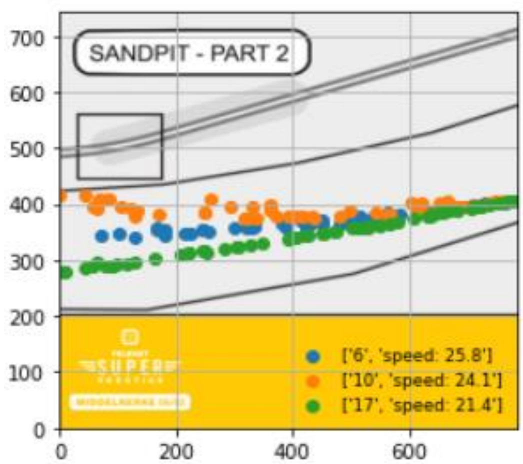

Figure 12: Riding lines analysed of top 3 riders in the sand pit zone of a pro cyclocross race.

6. Conclusions

In this paper, we performed an automatic sprint line detection on bird's eye view video footage shot by helicopters. These video frames are analyzed by a series of computer vision algorithms. Riders, road and broadcaster logos are detected on each frame. Riders are tracked across frames using a SORT object tracker. To visualize and summarize the gathered information, the separate frames are stitched together using a SIFT feature matcher that defines the geometric transformation that needs to be applied on a certain frame to geospatially stitch the frames into a final finishing straight.

The proposed methodology is very useful for further use in applications such as video referee assistance, sprint performance analysis or historical sprint similarity. In the future we aim to further extend the stitched sprint lines with rider GPS data and/or satellite data of the finishing straight to further understand the peloton's behaviour in the last few hundred meters.

This work was partly funded by the DAIQUIRI project, cofunded by imec, a research institute founded by the Flemish Government. Project partners are Ghent University, InTheRace, Arinti, Cronos, VideoHouse, NEP Belgium, and VRT, with project support from VLAIO under grant number HBC.2019.0053.

References

1. The all-important final meters: the sprint. (2021). Retrieved from https://cyclinglab.cc/en/the-all-important-final-meters-the-sprint/
2. Dilger, A., & Geyer, H. (2009). The dynamic of bicycle finals: A theoretical and empirical analysis of slipstreaming. Economic Analysis & Policy, 39, 429-442.
3. Lybbert, T. J., Lybbert, T. C., Smith, A., & Warren, S. (2012). Does the Red Flag Rule Induce Risk Taking in Sprint Finishes? Moral Hazard Crashes in Cycling's Grand Tours. Journal of Sports Economics, 13(6), 603–618. https://doi.org/10.1177/1527002511412077
4. Monier, E., Wilhelm, P., & Rückert, U. (2009). A computer vision based tracking system for indoor team sports. In *The fourth international conference on intelligent computing and information systems*.
5. Nie, X., Chen, S., & Hamid, R. (2021). A robust and efficient framework for sports-field registration. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (pp. 1936-1944).
6. Ghanem, B., Zhang, T., & Ahuja, N. (2012, March). Robust video registration applied to field-sports video analysis. In *IEEE*

International Conference on Acoustics, Speech, and Signal Processing (ICASSP) (Vol. 2).

7. Specification for organizers (2021). Retrieved from https://www.uci.org/docs/default-source/publications/2021/2021-uci-road-cahier-charges-orga-eng.pdf
8. Menaspa, P. (2015). Analysis of road sprint cycling performance. 2015
9. Cohen, C., Brunet, E., Roy, J., & Clanet, C. (2021). Physics of road cycling and the three jerseys problem. In *Journal of Fluid Mechanics, 914.*
10. Lowe, D. (2004). Distinctive image features from scale-invariant keypoints. In *International Journal of Computer Vision*, 60, 91-110.
11. Bewley, A., Ge, Z., Ott, L., Ramos, F. & Upcroft, B. (2016) Simple online and realtime tracking. In *IEEE International Conference on Image Processing (ICIP)*, 3464-3468.
12. Yuxin, W., Kirillov, A., Massa, F., Lo, WY.
13. & Girshick, R. (2019). Detectron2. Retrieved from https://github.com/facebookresearch/detectron2