SoccerNet Game State Reconstruction: End-to-End Athlete Tracking and Identification on a Minimap

Supplementary Material

A. Camera calibration

The estimated camera parameters follow the pinhole camera model augmented with one radial distortion coefficient that may be needed in wide camera shots. The camera parameters are estimated in four steps. First, as a global pre-processing step, the intersections of the pitch markings are computed to obtain both line-to-line and point-to-point correspondences between the image and the soccer pitch model. Then, depending on the visible parts of the soccer pitch in the image, different strategies are used to retrieve camera parameters. When there is a sufficient amount of pitch markings in an image, a homography mapping the image plane to the soccer pitch plane is estimated, then converted into pinhole camera parameters. Moreover, an optimization is conducted to determine one radial distortion coefficient given the curvature of the annotated polylines. For the frames that do not display enough pitch markings, the knowledge that each sequence is shot by a single camera is leveraged. As broadcast cameras are both zooming and rotating, only the camera position can be considered fixed. It is estimated as the median 3D position of the camera parameters estimated in the previous step. A similar version of the Two-Point PTZ algorithm is used to compute the focal length, pan and tilt parameters. Finally, as there are still some frames that can not be calibrated with sufficient accuracy, an industrial tool is used to compute the camera parameters of the missing frames.

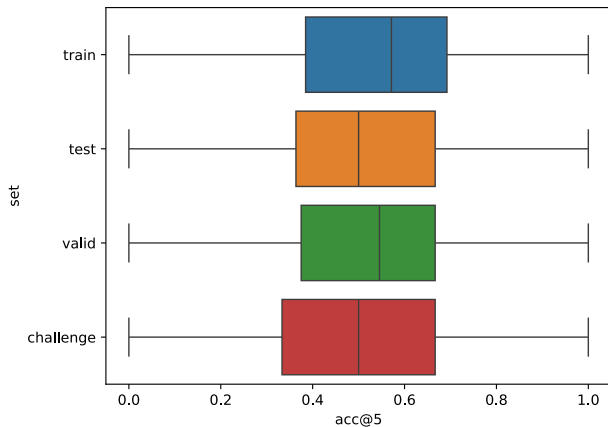Some examples of the pitch annotations used for the camera calibration can be found in Fig. S2.

Figure S1. Distribution of the `acc@5` metric for the different sets

B. GS-HOTA additional discussion

The HOTA authors introduced a "Classification-Aware HOTA" that shares similarities with our proposed GS-HOTA. However, the Classification-Aware HOTA is not suitable for evaluating game

Table S1. **GSR-Baseline on SoccerNet Tracking [20].** *FT Det* indicate an object detector fine-tuned on SoccerNet. GSR-B uses an out-of-the-box YOLOv8, not fine-tuned on soccer data.

Algorithm	FT Det	HOTA	DetA	AssA	MOTA
DeepSORT [88]		36.66	40.02	33.76	33.91
FairMOT [94]		43.91	46.32	41.78	50.70
ByteTrack [95]		47.23	44.49	50.26	31.74
GSR-B (ours)		**57.64**	**67.42**	49.42	**80.79**
FairMOT-ft [94]	✓	57.88	66.56	50.49	83.56
SNT23-Winners [56]	✓	**73.29**	**73.26**	**73.42**	**87.74**

state reconstruction for several reasons: first, it imposes a less rigid constraint on class predictions than Eq. (5), second, it is tailored for a single classification objective, and third, it necessitates to output one classification score for each potential class, all summing up to one, an unsuitable requirement for team affiliation and jersey number recognition.

C. Comparison with Standard MOT Methods.

We present the performance of the 'image tracking only' component of our baseline, which includes the detector, ReID, and tracking modules, to compare with existing SOTA Multi-Object Tracking (MOT) methods. To this end, we employ two well-established metrics: MOTA [5] and HOTA [54]. Results in Tab. S1 reveal our superior performance over methods using non-fine-tuned object detectors. Finally, a specialized soccer tracking method such as [56] highlights the potential for improvement in image-based tracking. This method relies on a strong object player detector fine-tuned on soccer data, and a heavy test-time fine-tuning of a ReID model to associate short tracklets into long tracks and achieve long-term tracking.

D. Annotation sample

An annotation sample in JSON format is illustrated in Fig. S3 for a single video.

Figure S2. **Pitch annotations.** Examples of pitch annotations.

```json
{
    "info":{
      "version":"1.1",
      "game_id":"11",
      "id":"200",
      "num_tracklets":"20",
      "action_position":"956196",
      "action_class":"Shots on target",
      "visibility":"visible",
      "game_time_start":"2 - 15:41",
      "game_time_stop":"2 - 16:11",
      "clip_start":"941000",
      "clip_stop":"971000",
      "name":"SNGS-200",
      "im_dir":"img1",
      "frame_rate":25,
      "seq_length":750,
      "im_ext":".jpg"
    },
    "images": [
      {
         "is_labeled":true,
         "image_id":"3200000001",
         "file_name":"000001.jpg",
         "height":1080,
         "width":1920,
         "has_labeled_person":true,
         "has_labeled_pitch":true,
         "has_labeled_camera":true,
         "ignore_regions_y":[],
         "ignore_regions_x":[]
      },
      // Additional images annotations...
     ],
    "annotations":[
       {
          "id":"3200000001",
          "image_id":"3200000001",
          "track_id":1,
          "supercategory":"object",
          "category_id":1,
          "attributes":{
             "role":"player",
             "jersey":"14",
             "team":"left"
          },
          "bbox_image":{
             "x":1020,
             "y":508,
             "x_center":1043.0,
             "y_center":557.5,
             "w":46,
             "h":99
          },
          "bbox_pitch":{
             "x_bottom_left":-29.17307773076183,
             "y_bottom_left":-13.960906317008366,
             "x_bottom_right":-28.399824812615115,
             "y_bottom_right":-14.278786952621587,
             "x_bottom_middle":-28.786446826184775,
             "y_bottom_middle":-14.119801608871501
          }
       },
       // Additional athletes annotations...
       ...
```

```
66        ...
67        {
68         "id":"3200000019",
69         "image_id":"3200000001",
70         "supercategory":"pitch",
71         "category_id":5,
72         "lines": {
73           "Side line top":[{"x":0.21, "y":0.34}, {"x":0.61, "y":0.39}, {"x":1.0, "y":0.43}],
74           "Side line left":[{"x":0.0, "y":0.45}, {"x":0.10, "y":0.39}, {"x":0.21, "y":0.34}],
75           "Small rect. left top":[{"x":0.07, "y":0.53}, {"x":0.01, "y":0.53}, {"x":0.01, "y":0.53}],
76           "Small rect. left main":[{"x":0.0, "y":0.54}, {"x":0.01, "y":0.54}, {"x":0.01, "y":0.53}],
77           "Big rect. left top":[{"x":0.04, "y":0.42}, {"x":0.23, "y":0.45}, {"x":0.41, "y":0.48}],
78           "Big rect. left main":[{"x":0.0, "y":0.81}, {"x":0.20, "y":0.65}, {"x":0.41, "y":0.48}],
79           "Circle left":[{"x":0.02, "y":0.79}, {"x":0.04, "y":0.79}, ...],
80         }
81       }
82      // Additional pitch annotations...
83     ]
84 }
```

Figure S3. Sample JSON annotation for one video of the SoccerNet-GSR dataset