A The Airport Dataset

The dataset consists of 12 sequences from the airport apron surveillance scenario, displayed and explained in Fig. 1. It contains sequences with large objects (*firetruck*, *rolling3*), small objects (*followme1*, *followme2*, *landing2*), non-rigid objects (*rolling1*), fast motion (*landing1*), slow motion (*push-back*), low framerate (*heli*), moving camera (*panning*), zooming camera (*bus*), and non-planar motion (*rolling2*).

All sequences have a resolution of 1440×1080 and a length of $\approx 10\,\mathrm{s}$ at a frame rate of $4\,\mathrm{Hz}$. The frame rate was selected as reasonable value in order to allow realtime motion segmentation without having too much change in appearance to establish correspondences between consecutive frames. Only the *heli* sequence has a resolution of 1280×960 at a non-steady frame rate between 1 and 2 Hz (typical for a network surveillance camera).

In every sequence, 8 to 16 frames $\mathbf{J}_t$ are manually annotated with a ground truth image $\mathbf{G}_t$ (Fig. 2). A pair of ground-truth-annotated frames allows to automatically categorize every established image correspondence as true/false inlier/outlier.

For convenience, we provide

- Matlab functions for automatic evaluation (Sec. A.1),
- the SIFT correspondences we established between the ground truth frames, as well as
- our segmentation results for RMSAM (sam3), J-Linkage (jl2) and GBS (gbs3).

A.1 Matlab functions for automatic evaluation

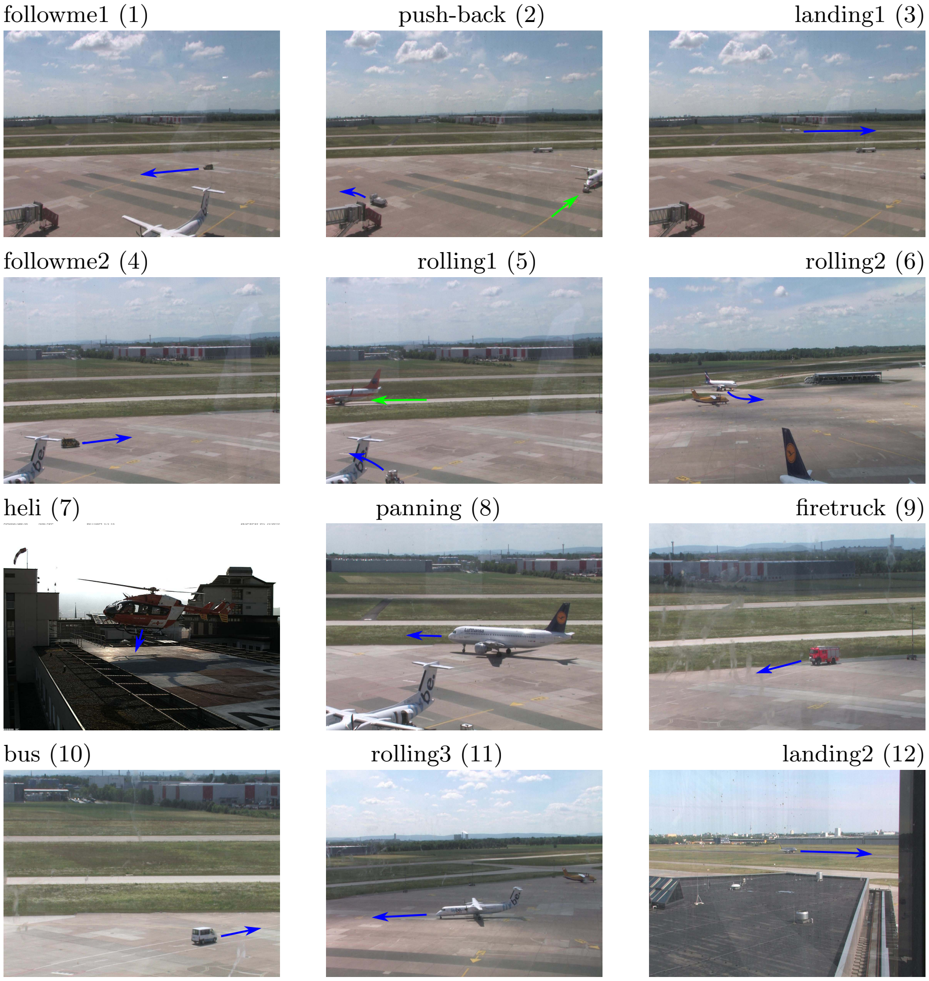
- `get_gt_from_images(corrfile, I1, I2, gtlabelfile)`
  Given correspondences from `corrfile` between images `I1` and `I2`, create the `gtlabelfile` containing a ground truth labeling of `corrfile`. With this function, the ground truth segments for own correspondences can be found.

-
  ```
  [op, or] = compare_labeling_with_gt(labelfile, ...
                                  gtlabelfile, n_gt_classes)
  ```
  Given the `labelfile`, the `gtlabelfile` and the number of ground truth classes, compute object precision `op` and recall `or`.
- `evaluate_jl_gbs_msam`
  The evaluation of the `hannover1_results` directory. By changing the prefixes in the `approach` cell, it should be easy to incorporate own segmentation results.

A.2 File Formats

For the following file formats, the space characters can be replaced by newline characters.

- Correspondences file to store point correspondences between two images. Each line consists of one correpondence string:
  $-1\ x_1\ y_1\ -1\ x_2\ y_2$
  e.g. `-1 50 50 -1 100 100`
- Label file to label correspondences between two images. The first segment has id 0 in the file. In matlab, this corresponds to segment 1. The outlier segment has the id -1 and 0 in matlab, respectively. The label vector $\boldsymbol{l} = (l_1, l_2, \ldots l_N)^T \in \mathbb{N}^N$ is encoded with the following string:
  1 labels 1 $N\ l_1\ l_2\ l_3\ \ldots\ l_N$
  e.g. `1 labels 1 5 1 0 0 0 1`



followme1 (1)

push-back (2)

landing1 (3)

followme2 (4)

rolling1 (5)

rolling2 (6)

heli (7)

panning (8)

firetruck (9)

bus (10)

rolling3 (11)

landing2 (12)

Fig. 1. Sequences from the airport dataset from top left to bottom right: *followme1* (the follow-me car is moving), *push-back* (the luggage cart and the airplane are moving), *landing1* (the airplane is landing on the runway), *followme2* (the follow-me car is moving), *rolling1* (the airplane on the taxi way is moving as well as the luggage car), *rolling2* (the distant airplane is rolling and turning). *heli* (the helicopter is landing), *panning* (the camera is panning while the airplane is rolling), *firetruck* (the fire truck is moving), *bus* (the bus is moving and the camera zooming in), *rolling3* (the airplane is rolling), *landing2* (the airplane is landing on the distant runway). The numbers in parentheses denote the sequence number.

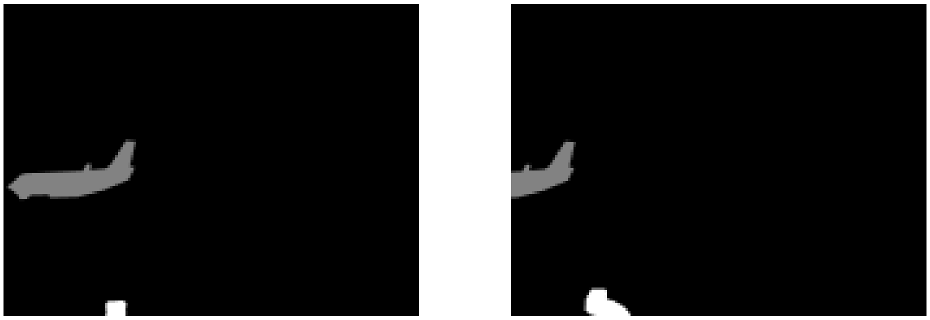


Fig. 2. Examples of the ground truth annotation for 3 motion segments of the *rolling1* sequence. Keypoint correspondences between same-colored regions are considered true, and from differently colored regions as false.