Joint Vanishing Point Extraction and Tracking (Supplementary Material)

Till Kroeger[1] Dengxin Dai[1] Luc Van Gool[1,2]

[1] Computer Vision Laboratory, D-ITET, ETH Zurich
[2] VISICS, ESAT/PSI, KU Leuven
{kroegert, vangool}@vision.ee.ethz.ch

Annotation of the VP detection and tracking dataset

We use a dataset which was published [1] for the task of registering videos to Structure from Motion (SfM) point clouds. The datasets consists of 4 sets of street-view videos, with 12 videos in each set, and a total of 48 videos. The 12 videos in each set were captured simultaneously from 12 cameras, which were rigidly mounted onto a van, in different parts of the town of Antwerp at a framerate of 10 fps. For each of the 4 sets a SfM point cloud, corresponding 2D feature locations, and precise 6-DoF camera poses for all views and frames are provided. Fig. 1 shows for one set (S01) the SfM point cloud (left), the camera set-up on the van (middle), and example images from the cameras (right).

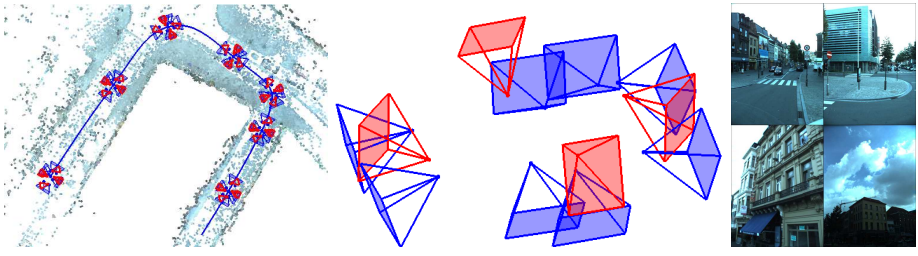

Fig. 1. Left: SfM point cloud, with van path over 301 frames and selected camera poses. Middle: Rigid camera set-up on the van. The red/blue coloring is irrelevant for our purposes. Right: Example images from 4 out of 12 cameras at the same time point. The figures were adapted from [1] with permission from the authors.

Since there exists no dataset for VP detection and tracking over time, we adapted this dataset by adding VP annotation. We chose this dataset because ground truth camera poses are available. This enables, firstly, semi-automatic VP annotation, as described in the following, and, secondly, the experiment in the paper (§ 4.3) for the inclusion of pose knowledge in VP extraction and tracking.

For our evaluation we need annotation of (possibly multiple) VPs in each frame, and VP identities across time. Since the dataset consists of a total of

978-1-4673-6964-0/15/$31.00 ©2015 IEEE

14448 (= $4 \times 12 \times 301$) video frames, manual VP annotation and linking over time seems infeasible.

We adopted a semi-automatic procedure by using the known camera orientations, loosely based on [2], and proceed as follows for each of the four sets:

1. In 12 videos with each 301 frames, we detect line segments (LSD line detector [3]) in each frame, convert them to homogeneous normalized image coordinates (using the known internal camera calibration), and compute the corresponding interpretation planes as described in § 2 of the paper.
2. Since the precise 6-DoF camera pose for every frame is known we rotate all interpretation planes into the common world frame.
3. For (on average) $\sim$ 1100 image line segments in each frame (3612 frames $= 12 \times 301$), we compute $\sim$ 4 000 000 interpretation planes. We keep the 50 000 interpretation planes corresponding to the longest line segments. To ensure good coverage of all frames, we keep at least 10 planes (corresponding to the longest line segments) from each frame.
4. In order to create VP exemplars we randomly sample (with replacement) 1000 pairs of intersection planes and compute the resulting vanishing direction for every pair (See § 2 of the paper).
5. For every interpretation plane we compute the consistency to all VP exemplars as: $|\alpha - 90|$, where α is the angle between plane normal and exemplar vanishing direction.
6. Having such VP exemplars and interpretation plane consistency scores, we perform J-Linkage clustering, as used in [2] for VP detection. This process results in a small set of dominant scene directions in the world coordinate system of one set of 12 videos. Small VP clusters (with less than 1 percent of all interpretation planes associated to it) are removed.
7. We decide which of the dominant scene directions is visible in a given frame based on the number of LSD line segments, extracted from this frame, which are consistent with it. We chose a high threshold to trigger the start of VP track and a lower threshold to check for continuation in following frames. Selecting two thresholds (one for starting, one for continuation) ensures that a VP track is only started when the VP detection is reliable, but enables continuation through noisier frames with weaker VP line support. This process is very similar to the high and low hysteresis thresholds used for canny edge detection. We randomly assign numeric ground truth IDs to VP tracks created in this way.
8. For each frame in each video, we rotate all visible vanishing directions back to the camera coordinate system, using the known camera pose, and store the direction.
9. The resulting VP tracks are manually inspected and occasional small errors (start/ending times and continuation of VP tracks) are corrected.

The whole process is inspired by single-frame VP detection in [2], but differs in two significant ways from it: Firstly, we jointly perform VP detection in multiple views and over time using known poses. Secondly, we reason with a Gaussian sphere parametrization, while [2] parametrizes the VPs in image space.

Example images of all 48 sequences can be seen in Fig. 2 and 3. Extracted line segments are overlaid. The line color indicates ground truth VP identity.

Possible Errors. Since we annotate semi-automatically using another VP detection algorithm we have to discuss possible errors introduced into the annotation. We assume error-free ground truth poses, and error-free internal camera calibration given in the dataset. We also assume a negligible (orientation) error in the LSD line detection. Thus, interpretation planes from all lines of all frames will be transferred with negligible error into the global world reference frame.

Our approach, based on [2], generally has a very good recall (VPs are rarely missed, and accurately localized), but may suffer from slightly worse precision (false positive VPs are possible). This is due to the fact that some line segments will be consistent with multiple VP hypotheses on the horizon line. After visual inspection we found that filtering by line support strength (step 6) removes all false positive VPs. However, the question of false positive is also partly a question of which level of detail in annotation is required. Since every set of two parallel lines (e.g. outline of roof of a house, bricks in a wall) defines a VP, we have to select a target VP sensitivity. We designed the sensitivity such that dominant VPs in the scene over multiple frames are captured, but short-lived VPs with weak support (e.g. from slanted roofs) are not captured anymore. This trade-off is selected such that the returned level of detail of VP tracks is best suited for most potential tasks (3D Reconstruction, Object Tracking, Autonomous Navigation) in this street-view scenario.

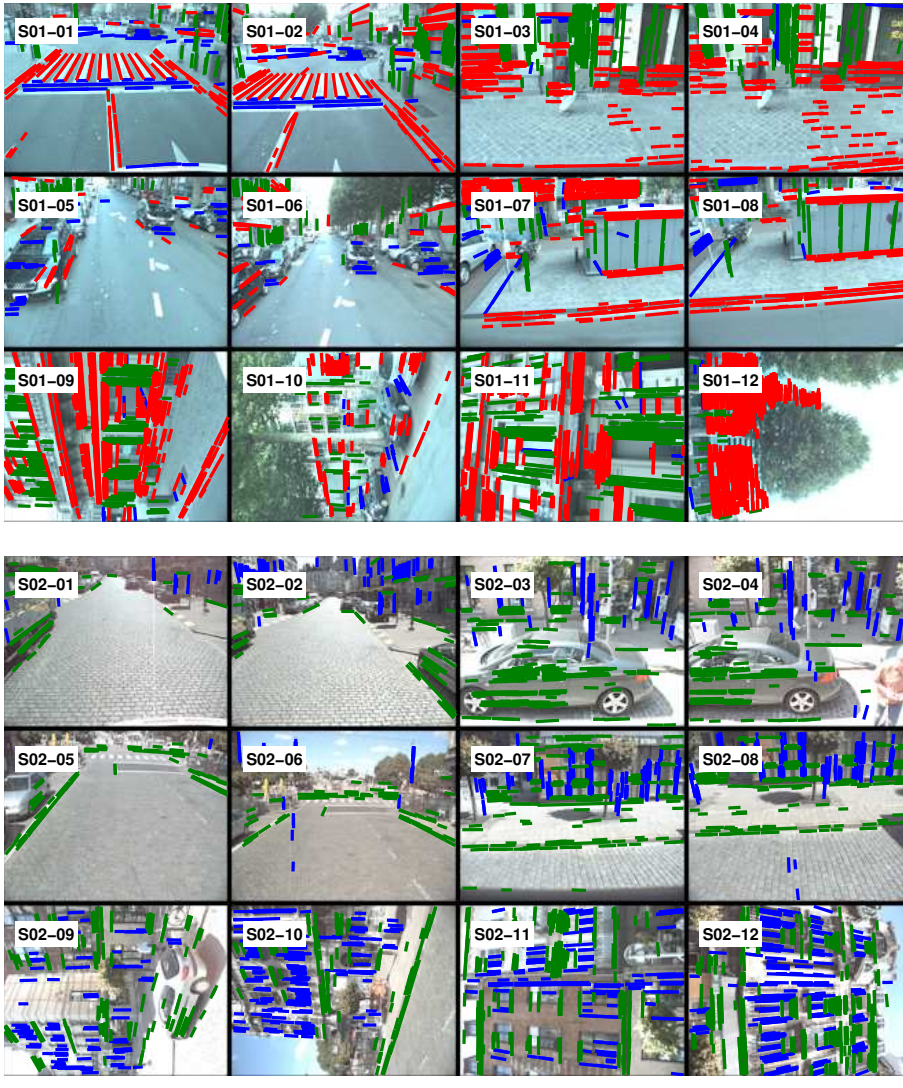

Fig. 2. All 12 views from video set 1 (top) and 2 (bottom) at a random frame with detected line segments. Line color indicates VP ground truth identity. Cameras 9:12 are rotated because they are mounted on the van in portrait mode.

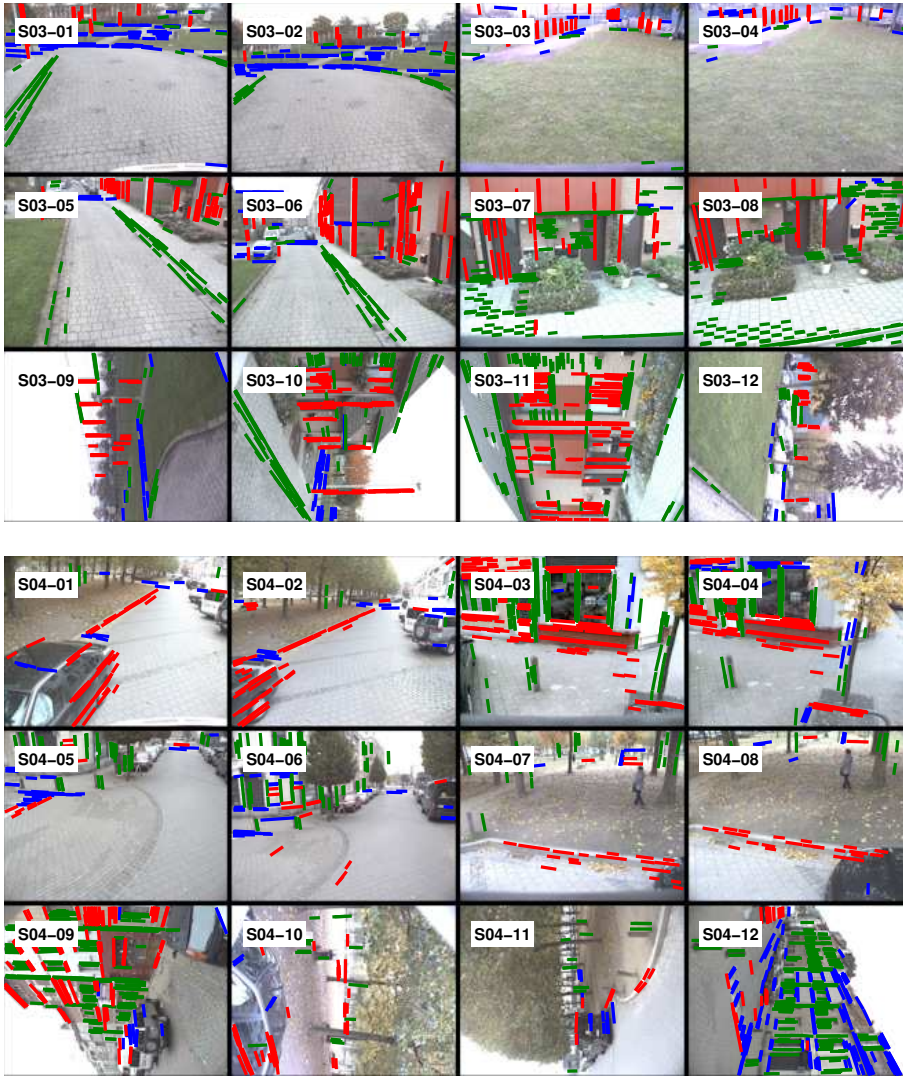

Fig. 3. All 12 views from video set 3 (top) and 4 (bottom) at a random frame with detected line segments. Line color indicates VP ground truth identity. Cameras 9:12 are rotated because they are mounted on the van in portrait mode.

Quantitative Evaluation on Street-View Dataset

Table 1 lists MOTA, Table 2 lists MOTP and Table 3 lists ID switches for all 48 sequences in 4 sets. Table 4 compares our method and the best baseline when knowledge about the precise camera orientation is incorporated. The best result in each row is printed in bold, the second best is underlined.

	T 0.6	T 0.6 f	T 0.5	T 0.5 f	T 0.4	T 0.4 f	T 0.3	T 0.3 f	LPSM	LPMF
S01-01	−0.18	−0.10	−0.13	−0.08	0.35	0.26	0.11	0.00	**0.59**	0.56
S01-02	−0.11	0.05	0.03	−0.07	0.47	0.42	0.27	0.05	0.58	**0.64**
S01-03	−0.33	−0.01	0.06	0.00	0.24	0.33	0.26	0.16	0.46	**0.51**
S01-04	−0.60	−0.22	0.02	0.03	0.18	0.28	0.31	0.23	0.38	**0.51**
S01-05	−0.63	−0.41	−0.19	−0.02	0.18	**0.19**	0.14	0.08	0.07	0.16
S01-06	0.02	0.10	0.08	−0.02	**0.51**	0.40	0.31	0.09	0.44	0.45
S01-07	−0.47	−0.03	0.09	0.06	0.38	0.48	0.41	0.21	0.55	**0.57**
S01-08	−0.79	−0.37	0.01	0.01	0.22	0.35	0.38	0.14	0.44	**0.53**
S01-09	0.13	0.43	0.54	0.39	0.63	0.65	0.64	0.48	0.65	**0.73**
S01-10	−0.22	0.07	0.10	0.08	0.43	0.50	0.36	0.20	0.49	**0.51**
S01-11	0.33	0.58	0.72	0.53	0.72	0.78	**0.83**	0.60	0.74	0.79
S01-12	−0.27	−0.10	−0.11	0.00	0.11	0.20	0.10	0.12	0.12	**0.26**
S02-01	−1.28	−0.88	−0.24	−0.33	−0.33	-0.11	**0.10**	−0.26	−0.68	−0.58
S02-02	−0.92	−0.61	−0.25	−0.08	−0.01	0.06	**0.10**	−0.04	−0.37	−0.45
S02-03	−1.67	−0.92	−0.44	−0.26	0.27	0.34	0.29	0.01	0.67	**0.83**
S02-04	−1.77	−0.95	−0.47	−0.26	0.17	0.29	0.25	−0.04	0.67	**0.78**
S02-05	−1.69	−0.96	−0.35	−0.09	0.00	**0.20**	0.17	−0.03	−0.10	0.04
S02-06	−0.70	−0.46	−0.16	−0.02	0.12	−0.01	0.14	0.06	0.33	**0.40**
S02-07	−1.21	−0.65	−0.30	−0.29	0.33	0.31	0.21	−0.02	0.62	**0.70**
S02-08	−1.43	−0.83	−0.44	−0.27	0.05	0.10	0.08	0.00	0.50	**0.57**
S02-09	−0.63	−0.19	0.18	0.06	0.38	0.46	0.49	0.18	0.63	**0.69**
S02-10	−0.52	−0.13	0.09	−0.02	0.38	0.42	0.42	0.18	0.51	**0.60**
S02-11	0.68	0.77	0.72	0.60	0.87	0.87	0.80	0.65	0.90	**0.92**
S02-12	0.05	0.30	0.39	0.18	0.61	0.57	0.51	0.33	0.75	**0.81**
S03-01	−0.09	−0.13	−0.12	0.04	**0.36**	0.25	0.16	0.08	0.24	0.17
S03-02	−0.01	0.04	0.00	−0.04	**0.46**	0.37	0.25	0.05	0.22	0.11
S03-03	−0.34	−0.29	−0.37	−0.20	0.16	0.15	−0.03	−0.04	0.37	**0.44**
S03-04	−0.56	−0.59	−0.43	−0.07	0.10	0.01	−0.01	0.03	0.37	**0.45**
S03-05	−0.34	−0.24	−0.12	−0.06	**0.20**	0.15	0.17	0.02	0.08	0.10
S03-06	−0.09	−0.08	0.01	−0.05	0.29	**0.30**	0.27	0.07	0.10	0.02
S03-07	−0.23	0.01	0.09	0.02	0.25	**0.33**	0.29	0.10	0.15	0.17
S03-08	−0.33	−0.10	−0.06	−0.04	0.18	**0.20**	0.15	0.03	0.16	0.14
S03-09	−0.59	−0.43	−0.36	−0.18	−0.04	**0.04**	−0.01	−0.02	0.03	−0.05
S03-10	−0.15	−0.03	0.04	−0.01	0.38	0.33	0.30	0.08	0.19	**0.43**
S03-11	−0.27	−0.17	−0.01	0.13	0.20	0.21	**0.31**	0.23	0.19	0.21
S03-12	−0.91	−0.79	−0.54	−0.25	−0.15	−0.18	-0.12	**−0.08**	−0.17	−0.16
S04-01	−7.69	−6.59	−4.05	−1.64	−3.76	−2.76	-1.51	**−0.43**	−2.73	−1.56
S04-02	−3.53	−3.07	−2.33	−1.20	−1.51	−1.19	−0.67	−0.36	-0.34	**0.20**
S04-03	−0.91	−0.49	−0.22	−0.11	0.18	0.34	0.30	0.21	0.41	**0.56**
S04-04	−0.99	−0.61	−0.32	−0.12	0.15	0.21	0.27	0.18	0.28	**0.48**
S04-05	−3.73	−2.55	−1.72	−0.36	−1.30	−0.83	−0.60	-0.06	−0.77	**0.04**
S04-06	−2.28	−1.68	−0.96	−0.34	−0.62	−0.35	0.13	0.04	−0.03	**0.42**
S04-07	−9.24	−6.66	−4.04	−1.28	−2.87	−1.82	-0.89	**−0.15**	−2.17	−1.18
S04-08	−17.58	−11.51	−6.74	-2.30	−5.44	−3.91	−2.35	**−0.39**	−5.32	−3.46
S04-09	−0.55	−0.10	0.20	0.35	0.46	0.68	0.64	0.55	0.62	**0.75**
S04-10	−27.53	−22.69	−15.62	−5.88	−9.84	−8.81	−5.38	**−1.06**	−6.69	-3.91
S04-11	−5.14	−3.53	−1.92	0.03	−1.52	−0.69	−0.07	**0.49**	−0.47	0.05
S04-12	0.00	0.29	0.54	0.42	0.67	0.71	**0.80**	0.54	0.73	0.79

Table 1. MOTA for all 48 sequences. **Best** and second best are marked.

	T 0.6	T 0.6 f	T 0.5	T 0.5 f	T 0.4	T 0.4 f	T 0.3	T 0.3 f	LPSM	LPMF
S01-01	2.30	2.28	2.18	1.89	2.29	2.27	2.09	**1.78**	2.19	2.28
S01-02	2.35	2.32	2.10	2.05	2.32	2.32	2.06	**2.00**	2.34	2.39
S01-03	2.41	2.42	2.60	2.34	2.40	2.42	2.62	2.27	2.25	**2.24**
S01-04	2.37	2.44	2.49	2.15	2.36	2.45	2.45	**2.01**	2.28	2.24
S01-05	2.49	2.44	2.39	2.10	2.45	2.44	2.34	**2.03**	2.32	2.28
S01-06	2.33	2.30	2.18	2.18	2.32	2.28	2.13	**2.10**	2.21	2.24
S01-07	2.32	2.08	2.14	2.08	2.29	2.03	2.05	1.91	1.91	**1.78**
S01-08	2.18	2.01	1.95	2.00	2.12	1.97	1.84	1.88	1.87	**1.81**
S01-09	1.97	2.06	2.05	1.99	**1.95**	2.04	2.03	1.98	2.13	2.11
S01-10	1.98	2.10	2.00	2.04	1.95	2.09	1.91	2.00	**1.82**	1.84
S01-11	1.52	1.48	1.60	1.84	1.50	1.47	1.59	1.80	1.50	**1.47**
S01-12	2.03	2.06	2.08	1.96	2.01	2.02	2.04	**1.95**	2.17	2.17
S02-01	1.75	1.92	2.25	1.68	1.71	1.87	2.28	1.78	2.24	**1.45**
S02-02	2.26	2.06	1.43	1.23	2.24	2.04	1.32	**0.92**	2.52	3.49
S02-03	2.19	2.32	2.52	2.79	2.14	2.25	2.44	2.84	1.99	**1.97**
S02-04	2.18	2.24	2.36	2.84	2.16	2.20	2.34	2.79	2.00	**2.00**
S02-05	1.52	1.37	1.26	1.28	1.44	1.30	**1.16**	1.41	1.96	2.05
S02-06	2.08	1.84	1.44	1.05	2.07	1.82	1.38	**0.98**	2.53	2.70
S02-07	2.41	2.46	2.53	2.55	2.42	2.41	2.53	2.50	2.38	**2.32**
S02-08	2.64	2.65	2.59	2.63	2.60	2.65	2.60	2.61	2.54	**2.47**
S02-09	2.61	2.60	2.52	2.54	2.62	2.59	2.51	**2.48**	2.87	2.87
S02-10	2.09	2.05	2.17	2.09	2.09	2.04	2.14	**2.00**	2.36	2.38
S02-11	1.81	1.77	1.76	2.06	1.80	1.76	**1.75**	2.04	1.78	1.76
S02-12	1.80	1.92	2.01	2.08	**1.79**	1.91	2.00	2.07	1.94	1.90
S03-01	1.85	1.88	1.62	**1.24**	1.83	1.88	1.58	1.31	2.40	2.47
S03-02	1.85	1.78	1.93	**1.74**	1.81	1.76	1.90	1.82	2.31	2.51
S03-03	2.62	2.65	2.49	**2.31**	2.62	2.67	2.58	2.33	2.32	2.38
S03-04	2.49	2.45	2.48	1.37	2.42	2.39	2.29	**1.17**	2.17	2.15
S03-05	1.56	1.48	1.25	0.89	1.42	1.39	1.08	**0.78**	2.26	2.28
S03-06	1.50	1.48	1.64	1.71	**1.46**	1.46	1.56	1.70	1.97	2.11
S03-07	2.50	2.36	2.24	2.30	2.49	2.32	**2.18**	2.28	2.36	2.30
S03-08	2.24	2.32	2.40	2.65	**2.19**	2.29	2.34	2.45	2.35	2.32
S03-09	2.27	2.44	2.52	2.35	2.18	2.37	2.50	2.44	2.06	**2.06**
S03-10	1.60	1.55	1.59	1.91	1.49	1.45	1.49	1.88	1.42	**1.41**
S03-11	1.40	1.50	1.61	1.58	1.30	1.35	1.51	1.53	1.21	**0.90**
S03-12	2.34	2.21	2.30	2.90	2.24	2.06	2.18	2.48	2.20	**1.97**
S04-01	1.86	2.15	3.12	NaN	**1.77**	2.07	3.31	NaN	2.04	2.01
S04-02	1.44	1.64	2.03	2.79	**1.39**	1.61	2.04	2.81	1.51	1.60
S04-03	2.14	2.28	2.45	2.21	2.14	2.29	2.44	2.18	2.05	**2.03**
S04-04	2.13	2.20	2.32	2.20	2.11	2.19	2.29	2.14	1.99	**1.97**
S04-05	2.43	2.44	2.42	2.66	2.43	2.43	2.43	2.69	**2.02**	2.08
S04-06	1.96	1.81	1.98	2.61	1.91	**1.74**	1.89	2.53	1.99	2.02
S04-07	2.78	2.81	3.04	2.97	2.84	2.94	3.07	2.90	**2.57**	2.67
S04-08	2.84	2.47	2.47	2.44	2.83	2.44	2.42	**2.33**	2.53	2.60
S04-09	1.52	1.49	**1.46**	1.53	1.52	1.49	1.46	1.52	1.78	1.84
S04-10	2.52	2.43	**2.32**	2.60	2.52	2.43	2.32	2.97	2.55	2.80
S04-11	1.67	1.62	1.58	1.49	1.66	1.62	1.59	**1.44**	1.89	1.82
S04-12	1.55	1.62	1.71	1.86	1.54	1.61	1.72	1.82	**1.49**	1.55

Table 2. MOTP (in degrees) for all 48 sequences. **Best** and second best are marked.

	T 0.6	T 0.6 f	T 0.5	T 0.5 f	T 0.4	T 0.4 f	T 0.3	T 0.3 f	LPSM	LPMF
S01-01	6	13	12	2	5	8	4	1	3	**0**
S01-02	12	5	7	6	6	3	3	1	4	**0**
S01-03	21	8	8	11	13	6	1	4	4	**0**
S01-04	12	6	20	11	4	2	4	5	3	**0**
S01-05	18	10	9	3	9	3	3	**0**	4	1
S01-06	16	7	12	7	6	3	4	3	2	**1**
S01-07	25	23	13	16	10	9	2	6	4	**0**
S01-08	15	15	13	8	6	6	4	2	5	**0**
S01-09	7	8	4	10	3	2	2	4	0	**0**
S01-10	13	10	11	7	5	7	2	2	4	**0**
S01-11	17	11	8	15	15	6	6	5	3	**0**
S01-12	8	8	9	4	3	3	2	3	**1**	3
S02-01	5	6	5	0	3	3	2	0	0	**0**
S02-02	9	8	1	0	5	3	0	0	0	**0**
S02-03	25	26	17	10	9	7	4	3	3	**0**
S02-04	21	27	17	5	8	8	5	0	3	**0**
S02-05	6	8	1	0	2	3	0	0	0	**0**
S02-06	7	6	3	0	6	3	2	0	0	**0**
S02-07	17	14	12	4	8	2	3	2	6	**1**
S02-08	23	12	11	9	9	9	5	3	11	**0**
S02-09	14	15	5	19	7	8	2	3	1	**0**
S02-10	7	7	8	7	4	5	4	4	**0**	1
S02-11	6	7	8	11	3	1	6	2	0	**0**
S02-12	5	6	5	10	1	3	0	3	2	**0**
S03-01	6	7	2	0	2	4	0	0	2	**0**
S03-02	16	10	8	2	7	4	1	**0**	5	1
S03-03	20	13	9	**0**	7	3	1	0	3	1
S03-04	18	11	5	0	5	1	1	0	5	**0**
S03-05	10	2	6	1	3	1	2	0	1	**0**
S03-06	6	4	9	5	3	4	2	2	2	**0**
S03-07	11	16	7	3	3	7	0	1	4	**0**
S03-08	6	11	3	5	2	8	3	1	5	**0**
S03-09	9	12	6	1	1	1	3	0	3	**0**
S03-10	12	9	7	5	2	1	3	1	2	**0**
S03-11	6	6	3	1	2	1	0	1	0	**0**
S03-12	10	8	4	0	0	0	0	0	1	**0**
S04-01	2	3	3	0	0	2	0	0	0	**0**
S04-02	1	2	2	1	0	1	1	0	0	**0**
S04-03	12	12	13	12	7	6	7	5	1	**1**
S04-04	14	10	8	15	7	5	4	5	2	**2**
S04-05	5	5	2	1	3	2	1	1	1	**0**
S04-06	8	4	11	3	2	0	3	0	1	**0**
S04-07	0	2	2	0	0	0	1	0	1	**0**
S04-08	5	1	1	2	2	1	**0**	0	0	1
S04-09	7	**0**	1	3	7	0	1	1	2	2
S04-10	0	0	0	1	0	0	0	0	0	**0**
S04-11	3	1	2	1	3	1	1	0	0	**0**
S04-12	1	5	1	6	1	3	1	1	**0**	1

Table 3. ID switches for all 48 sequences. **Best** and second best are marked.

	T 0.4 f	T 0.f f kp	LPMF	LPMF kp
S01-01	0.29	**0.46**	0.46	0.43
S01-02	0.26	0.41	0.51	**0.55**
S01-03	0.23	0.17	0.43	**0.49**
S01-04	0.20	0.29	**0.52**	0.52
S01-05	0.19	**0.47**	0.05	0.05
S01-06	0.30	**0.50**	0.30	0.38
S01-07	0.34	0.45	0.55	**0.58**
S01-08	0.30	0.32	**0.75**	0.65
S01-09	0.54	**0.73**	0.72	0.72
S01-10	0.31	0.38	**0.48**	0.28
S01-11	0.68	0.75	0.72	**0.77**
S01-12	0.32	**0.36**	0.25	0.16
S02-01	0.08	**0.12**	−0.61	−0.63
S02-02	**0.15**	0.02	−0.39	−0.45
S02-03	0.26	0.27	**0.77**	0.77
S02-04	0.23	0.13	**0.79**	0.77
S02-05	0.36	**0.41**	−0.15	0.08
S02-06	0.20	0.17	**0.31**	0.29
S02-07	0.50	0.29	**0.69**	0.65
S02-08	0.24	0.32	**0.53**	0.42
S02-09	0.52	0.61	0.71	**0.77**
S02-10	0.47	0.53	0.58	**0.73**
S02-11	0.79	0.85	**0.97**	0.94
S02-12	0.56	0.65	0.79	**0.82**
S03-01	0.21	**0.27**	−0.14	−0.07
S03-02	0.37	**0.39**	0.22	0.05
S03-03	0.13	0.13	0.48	**0.52**
S03-04	0.08	0.04	0.46	**0.52**
S03-05	0.16	**0.22**	0.09	0.11
S03-06	0.26	**0.40**	0.00	−0.06
S03-07	0.20	**0.37**	0.20	0.22
S03-08	0.22	**0.28**	0.18	0.22
S03-09	0.08	**0.14**	0.04	−0.01
S03-10	0.30	0.36	**0.47**	0.38
S03-11	0.14	0.36	**0.40**	0.38
S03-12	0.02	−0.04	−0.09	**0.04**
S04-01	−1.15	**−0.54**	-0.62	−1.69
S04-02	−0.32	−0.36	**0.48**	0.32
S04-03	0.43	0.57	**0.67**	0.57
S04-04	0.40	0.40	**0.55**	0.51
S04-05	−0.50	-0.19	**−0.12**	−0.23
S04-06	0.17	0.46	0.03	**0.89**
S04-07	−0.17	−0.08	0.08	**0.42**
S04-08	**−0.71**	−1.14	-0.86	−1.00
S04-09	0.63	0.73	0.73	**0.76**
S04-10	−1.50	−2.25	-1.25	**−0.25**
S04-11	0.62	**0.69**	0.50	0.25
S04-12	0.67	**0.90**	0.79	0.76

Table 4. MOTA for all 48 sequences. **Best** and second best are marked.

Vanishing Point Detection, Precision and Recall

We formulated the problem as tracking problem, and therefore choose to evaluate with MOTA, which incorporates an equally weighted missed detection and false positive count. But if the problem is regarded as a detection task, and temporal continuity ignored, the VP tracks can be split into frame-wise VP detections. This enables easier evaluation in terms of detection precision and recall.

We extracted VP tracks as described in § 4 of the paper and split the tracks into independent VP detections in each frame. The proposed method is on par with the best performing baselines as shown by the precision and recall curve in Fig. 4

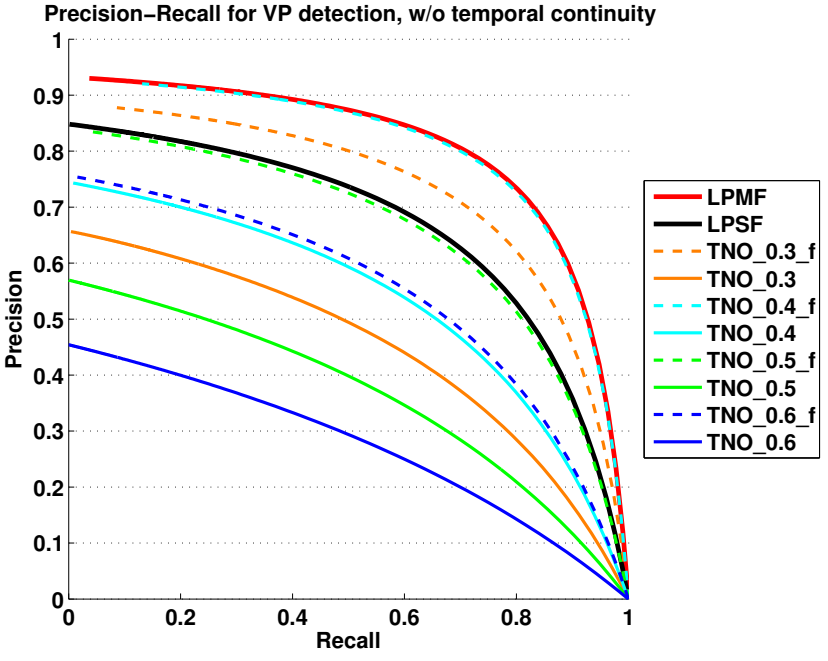

Fig. 4. Precision and recall for vanishing point detection with the proposed method. For this experiment we ignored temporal continuity of extracted vanishing point tracks, and regarded the problem as independent detection tasks in each frame.

References

1. T. Kroeger and L. Van Gool. Video Registration to SfM Models. In *ECCV*, 2014.
2. J.-P. Tardif. Non-Iterative Approach for Fast and Accurate Vanishing Point Detection. In *ICCV*, 2009.
3. R. G. von Gioi, J. Jakubowicz, J.-M. Morel, and G. Randall. LSD: a Line Segment Detector. *IPOL*, 2012.