

Table 1 Computational times (seconds) of each stage of our method on DAVIS 2016.

Sequences	Number of frames	Preprocessing	Graph-cut based video object segmentation				Total time	Time per frame
		Bilateral grid construction	Reliable and dynamic appearance modeling	Graph construction	Graph-cut optimization	Grid to masks		
bear	82	5.14	7.79	5.60	0.58	0.58	24.30	0.30
blackswan	50	3.04	5.03	5.37	0.75	0.19	16.06	0.32
bmx-bumps	90	5.37	12.54	6.17	3.88	0.34	30.97	0.34
bmx-trees	80	3.59	10.38	7.97	4.97	0.38	29.44	0.37
boat	75	3.31	14.84	4.77	2.45	0.31	27.83	0.37
breakdance	84	4.85	19.91	7.40	10.79	0.40	48.29	0.57
breakdance-flare	71	3.57	8.75	5.52	2.80	0.32	23.38	0.33
bus	80	3.44	13.19	6.29	7.33	0.34	32.98	0.41
camel	90	5.71	9.70	6.43	0.99	0.36	25.50	0.28
car-roundabout	75	3.98	13.57	7.35	4.55	0.61	33.07	0.44
car-shadow	40	2.66	5.51	4.38	2.98	0.14	17.23	0.43
car-turn	80	3.39	8.08	5.92	1.00	0.43	22.52	0.28
cows	104	6.63	19.32	7.56	1.83	0.55	39.09	0.38
dance-jump	60	3.23	5.74	5.68	1.20	0.32	18.51	0.31
dance-twirl	90	5.87	19.74	8.75	17.89	1.19	57.11	0.63
dog	60	3.26	6.22	4.40	0.52	0.27	21.25	0.35
dog-agility	25	2.68	5.14	5.19	1.06	0.16	16.00	0.64
drift-chicane	52	2.93	3.71	4.36	0.43	0.31	13.80	0.27
drift-straight	50	3.00	9.02	6.18	12.65	0.21	32.58	0.65
drift-turn	64	5.75	15.60	7.99	5.60	0.91	39.15	0.61
elephant	80	5.35	8.44	6.61	0.57	0.79	25.73	0.32
flamingo	80	3.63	6.91	5.48	1.21	0.30	19.96	0.25
goat	90	7.54	13.71	8.07	1.87	0.66	39.90	0.44
hike	80	3.52	12.37	5.16	0.70	0.53	24.64	0.31
hockey	75	4.59	9.84	5.22	1.99	0.27	23.97	0.32
horsejump-high	50	4.29	7.91	6.30	2.55	0.29	24.28	0.49
horsejump-low	60	6.56	12.36	7.52	6.17	0.38	37.23	0.62
kite-surf	50	2.92	4.50	4.53	1.04	0.19	14.90	0.30
kite-walk	80	3.29	6.13	4.62	1.07	0.38	17.62	0.22
libby	49	3.08	4.30	5.27	0.83	0.19	15.24	0.31
lucia	70	3.44	7.36	8.15	1.85	0.29	23.64	0.34
mallard-fly	70	3.57	6.54	8.52	1.40	0.69	22.79	0.33
mallard-water	80	3.59	6.87	6.71	1.29	0.47	22.17	0.28
motocross-bumps	60	3.09	7.37	5.06	3.66	0.23	21.17	0.35
motocross-jump	40	2.81	6.02	4.73	4.43	0.17	19.72	0.49
motorbike	43	2.91	7.06	4.79	4.45	0.16	20.92	0.49
paragliding	70	3.19	5.33	4.66	0.41	0.27	15.95	0.23
paragliding-launch	80	3.34	5.36	5.25	1.01	0.32	17.86	0.22

Sequences	Number of frames	Preprocessing	Graph-cut based video object segmentation				Total time	Time per frame
		Bilateral grid construction	Reliable and dynamic appearance modeling	Graph construction	Graph-cut optimization	Grid to masks		
parkour	100	6.08	8.89	6.32	2.50	0.41	26.62	0.27
rhino	90	6.60	9.73	7.26	0.78	1.25	29.46	0.33
rollerblade	35	2.79	8.12	5.13	2.84	0.18	20.66	0.59
scooter-black	43	4.41	15.04	7.28	6.24	0.21	35.89	0.83
scooter-gray	75	4.70	12.70	7.08	12.72	0.41	40.94	0.55
soapbox	99	6.84	11.90	7.55	6.20	0.64	37.10	0.37
soccerball	48	2.06	20.54	3.75	0.53	0.20	28.21	0.59
stroller	91	4.87	13.04	6.03	6.15	0.51	35.31	0.39
surf	55	2.98	5.99	4.55	1.49	0.21	16.95	0.31
swing	60	4.68	7.62	6.13	1.50	0.26	22.10	0.37
tennis	70	3.49	6.69	5.71	1.93	0.70	20.21	0.29
train	80	3.53	18.67	7.02	11.25	0.32	41.12	0.51
Average							0.40 ± 0.01	

Table 2 Computational times (seconds) of each stage of our method on YouTube-Objects.

Categories	Sequences	Number of frames	Preprocessing	Graph-cut based video object segmentation				Total time	Time per frame
			Bilateral grid construction	Reliable and dynamic appearance modeling	Graph construction	Graph-cut optimization	Grid to masks		
Aeroplane	0001	61	2.70	2.91	2.44	0.66	0.13	10.10	0.17
	0002	90	3.06	5.00	2.87	1.61	0.21	14.19	0.16
	0010	124	3.19	5.85	3.41	4.44	0.25	18.83	0.15
	0011	15	2.44	1.65	2.23	0.16	0.03	7.74	0.52
	0012	401	4.21	4.84	3.53	1.63	0.62	17.48	0.04
Bird	0001	201	3.69	4.32	2.85	1.21	0.46	14.42	0.07
	0007	201	5.89	8.21	6.85	5.30	1.04	30.99	0.15
	0010	201	4.01	13.38	4.44	2.82	0.54	27.24	0.14
	0011	201	4.60	3.66	2.94	0.88	0.48	14.55	0.07
	0012	201	7.05	6.99	3.95	2.53	2.90	26.68	0.13
	0014	201	6.56	9.58	5.24	10.84	1.10	36.55	0.18
Boat	0001	201	5.90	7.62	5.54	4.68	0.61	26.56	0.13
	0003	201	5.94	4.86	2.76	0.75	0.47	17.09	0.09
	0005	201	7.16	4.18	2.42	0.45	0.89	18.70	0.09
	0007	201	5.80	15.52	5.28	21.02	0.72	50.95	0.25
	0008	116	3.41	7.35	3.61	7.93	0.31	24.46	0.21
	0009	66	2.63	13.74	5.63	2.55	0.28	26.44	0.40
	0010	156	5.03	4.00	2.85	1.53	0.37	15.14	0.10
	0011	105	2.93	5.42	3.17	2.90	0.30	16.21	0.15
	0012	186	5.48	8.09	4.96	17.88	0.63	40.39	0.22
	0014	9	2.11	9.86	1.99	0.38	0.03	15.32	1.70
	0015	201	4.50	4.74	2.84	3.18	0.50	17.58	0.09
	0016	109	3.54	4.71	3.60	7.89	0.26	21.28	0.20
	0017	201	4.64	3.94	2.83	2.27	0.52	16.10	0.08
Cat	0001	100	2.43	10.23	7.55	36.73	0.35	58.83	0.59
	0004	313	2.88	7.12	4.88	3.93	0.25	20.49	0.07
	0006	401	8.42	15.68	6.28	8.21	1.29	43.63	0.11
	0008	246	4.38	11.14	4.60	36.44	0.68	59.58	0.24
	0010	401	7.19	4.83	3.11	3.32	1.00	22.19	0.06
	0011	401	5.32	10.96	7.05	8.28	1.50	37.72	0.09
	0012	401	6.84	4.75	4.83	4.00	1.69	26.17	0.07
	0013	401	7.29	9.26	4.68	2.44	1.01	28.10	0.07
	0014	236	4.28	13.39	5.96	25.32	0.69	52.12	0.22
	0015	239	4.93	6.68	3.81	2.65	0.70	21.44	0.09
	0016	244	4.59	12.48	6.69	8.06	0.75	35.21	0.14
	0017	125	3.65	16.41	8.42	26.40	0.42	57.60	0.46
	0018	53	3.11	3.99	2.06	0.30	0.13	11.14	0.21
	0020	261	3.24	6.28	3.62	8.19	0.60	25.45	0.10

Categories	Sequences	Number of frames	Preprocessing	Graph-cut based video object segmentation				Total time	Time per frame
			Bilateral grid construction	Reliable and dynamic appearance modeling	Graph construction	Graph-cut optimization	Grid to masks		
Car	0002	93	2.65	11.61	4.49	10.31	0.25	30.76	0.33
	0003	36	2.98	3.16	2.52	0.51	0.10	11.69	0.32
	0004	71	4.05	5.78	4.65	14.96	0.35	32.04	0.45
	0009	134	3.73	3.66	2.70	1.78	0.36	14.75	0.11
Cow	0003	201	3.55	4.53	3.30	2.46	0.40	15.92	0.08
	0005	201	4.21	8.94	6.11	4.82	0.63	27.08	0.13
	0006	201	4.11	2.98	4.20	15.84	0.56	29.90	0.15
	0009	175	6.00	9.93	6.82	4.57	0.57	31.89	0.18
	0011	170	5.67	10.23	5.62	6.44	0.46	30.76	0.18
	0014	9	2.10	2.26	1.91	0.05	0.07	7.51	0.83
	0015	10	2.07	2.25	1.83	0.13	0.03	7.34	0.73
	0016	201	7.75	7.17	5.90	2.87	1.03	28.99	0.14
	0022	141	3.98	6.81	3.81	5.99	0.29	24.83	0.18
Dog	0003	58	2.94	9.27	5.52	14.59	0.19	33.99	0.59
	0005	72	2.78	3.67	3.05	0.85	0.18	11.85	0.16
	0006	201	4.09	12.13	7.48	5.18	0.64	32.65	0.16
	0008	160	3.57	14.22	6.35	13.26	0.53	41.28	0.26
	0010	201	3.89	8.76	5.25	5.02	0.59	25.75	0.13
	0012	201	4.25	10.40	6.85	5.01	0.62	29.67	0.15
	0013	201	7.72	17.63	7.98	3.77	0.78	41.67	0.21
	0020	201	3.98	11.23	5.55	4.02	0.65	28.31	0.14
	0021	153	3.47	3.13	2.53	1.19	0.37	12.32	0.08
	0022	133	4.98	13.07	8.66	13.13	0.66	41.99	0.32
	0023	201	3.90	16.35	7.10	10.15	0.76	43.06	0.21
	0025	201	4.02	4.60	3.17	0.72	0.49	15.20	0.08
	0026	150	3.30	2.66	2.54	1.08	0.37	11.54	0.08
	0027	82	3.64	2.95	2.83	0.40	0.21	11.28	0.14
	0028	201	8.04	15.38	9.27	7.95	1.91	46.98	0.23
	0031	201	3.76	5.09	3.51	2.84	0.57	17.80	0.09
	0034	201	5.76	6.00	2.92	1.47	0.48	18.47	0.09
	0035	201	3.85	7.02	3.13	5.62	0.49	19.10	0.10
	0036	90	3.15	4.23	3.08	1.03	0.28	13.21	0.15
Horse	0009	401	5.92	4.96	7.65	0.98	1.19	23.74	0.06
	0011	289	10.62	19.12	9.97	15.65	1.33	60.02	0.21
	0012	119	3.44	5.28	4.85	1.12	0.32	16.85	0.14
	0014	151	7.14	9.45	5.94	2.31	0.40	29.14	0.19
	0024	157	6.45	7.30	5.89	0.68	0.34	23.63	0.15
	0025	61	2.81	3.97	4.02	0.46	0.11	12.78	0.21
	0026	201	3.92	6.20	3.64	6.40	0.54	22.84	0.11

Categories	Sequences	Number of frames	Preprocessing	Graph-cut based video object segmentation				Total time	Time per frame
			Bilateral grid construction	Reliable and dynamic appearance modeling	Graph construction	Graph-cut optimization	Grid to masks		
Motorbike	0001	19	2.23	2.77	2.89	0.54	0.20	9.71	0.51
	0003	45	2.61	7.38	3.66	1.38	0.14	16.45	0.37
	0006	94	2.78	3.87	2.34	0.66	0.22	11.23	0.12
	0009	201	5.87	10.41	5.52	12.87	0.54	37.96	0.19
	0013	201	4.02	9.96	5.05	47.24	0.60	69.20	0.34
	0014	108	3.18	7.23	4.03	3.83	0.37	20.25	0.19
Train	0001	201	3.91	7.01	4.62	4.06	0.60	22.33	0.11
	0003	83	2.75	6.45	2.72	1.15	0.22	14.68	0.18
	0008	201	3.99	12.00	5.18	10.69	0.67	34.93	0.17
	0024	56	2.76	12.15	3.69	5.92	0.16	26.02	0.46
	0025	201	3.74	9.97	5.75	3.35	0.77	26.56	0.13
Average								0.19 ± 0.03	

Table 3 Computational times (seconds) of each stage of our method on SegTrack v2.

Sequences	Number of frames	Preprocessing	Graph-cut based video object segmentation				Total time	Time per frame
		Bilateral grid construction	Reliable and dynamic appearance modeling	Graph construction	Graph-cut optimization	Grid to masks		
bird_of_paradise	98	2.48	3.06	3.36	0.56	0.22	10.90	0.11
birdfall	30	1.77	2.81	3.63	0.22	0.04	9.36	0.31
bm_x_1	36	1.98	5.23	3.03	0.30	0.09	11.58	0.32
bm_x_2	36	1.97	4.63	3.06	0.60	0.09	11.32	0.31
cheetah_1	29	1.78	2.96	2.96	0.15	0.03	8.74	0.30
cheetah_2	29	1.74	2.67	3.08	0.12	0.04	8.54	0.29
drift_1	74	2.31	10.42	4.27	3.14	0.19	21.50	0.29
drift_2	74	2.47	10.28	4.10	4.71	0.20	22.93	0.31
frog	279	3.05	4.65	3.01	0.17	0.42	12.71	0.05
girl	21	1.78	1.34	3.20	0.18	0.07	7.54	0.36
hummingbird_1	29	1.92	2.50	3.06	0.40	0.06	8.90	0.31
hummingbird_2	29	1.96	1.96	3.30	0.45	0.15	8.77	0.30
monkey	31	1.93	2.11	3.73	0.57	0.06	9.32	0.30
monkeydog_1	71	1.90	2.33	3.34	0.13	0.10	8.74	0.12
monkeydog_2	71	1.88	3.77	3.14	0.23	0.08	9.99	0.14
parachute	51	2.09	2.62	2.87	0.08	0.10	8.71	0.17
penguin_1	42	1.90	4.12	3.33	0.57	0.06	10.86	0.26
penguin_2	42	1.93	4.56	3.24	0.48	0.06	11.17	0.27
penguin_3	42	1.83	3.89	3.18	0.45	0.05	10.26	0.24
penguin_4	42	2.02	3.57	3.36	0.38	0.05	10.27	0.24
penguin_5	42	1.90	3.56	3.13	0.41	0.05	9.92	0.24
penguin_6	42	1.81	3.92	3.17	0.47	0.11	10.38	0.25
soldier	32	1.87	2.17	3.26	0.44	0.06	8.76	0.27
worm	244	2.89	4.28	3.36	0.35	0.36	12.59	0.05
Average							0.24 ± 0.01	