Defocus-based Direct Visual Servoing - Addendum

Guillaume Caron

Defocus-based Direct Visual Servoing – Addendum

Guillaume Caron

I. Addendum

This document is meant to serve as an addendum to [1]. The main purpose of this addendum is to provide supplementary results, namely a statistical study of the convergence domain of the Defocus-based Direct Visual Servoing [1] (DDVS). It is conducted by starting the visual servoing for a various set of initial poses, which values of the six degrees of freedom (DoF) are randomly drawn (Sec. II), to a single desired pose. Section III reports DDVS results and Section IV reports the results of Photometric Gaussian Mixture-based Visual Servoing [2] (PGM VS) from the same initial poses. A comparison of these results concludes the addendum.

II. Experimental protocol

Following a similar methodology than other direct visual servoing [3], [4], [2], a set of initial poses is computed with respect to a desired pose. DDVS is then run with the camera at the end-effector of a robot arm (Universal Robot 10) from each of these initial poses to check if it converges, or not, to the desired pose.

In this addendum, the initial poses are defined on 3 concentric spheres centered at the desired pose. Their radii are: 20 cm, 40 cm, 60 cm. These values were chosen so that they encompass the previous largest reported convergence domains of DVS methods [5], [2], [6], [7], *i.e.* 30 cm, as well as the new capabilities claimed by DDVS [1], *i.e.* up to 90 cm. As the latter large distance is only possible within a very small part of the robot workspace, the largest considered sphere radius is 60 cm. Even limiting to 60 cm, only a subpart of the robot workspace is available, although of a much larger volume than the sphere of 90 cm radius.

Then, for each sphere, 23 initial positions are randomly generated with uniform probability density (ViSP library [8] implementation of [9]). On each of these 69 positions, the rotation of the camera around its horizontal and vertical axes is computed to ensure that the content of images acquired at initial poses overlaps the one of the desired image (Fig. 4). Then, the rotation around the optical axis is randomly generated in the interval [-15 degrees, 15 degrees].

The 69 obtained poses are reported in Table I where images are captured and shown in Figures 1, 2 and 3. Depending on the robot arm configuration, the scene illumination is not constant (*e.g.* shadowed as in Fig. 1f or 1j).

Guillaume Caron is with CNRS-AIST Joint Robotics Laboratory, IRL, National Institute of AIST, Tsukuba, Japan and with Université de Picardie Jules Verne, MIS laboratory, Amiens, France guillaume.caron@u-picardie.fr

III. DDVS convergence domain

In detail, from the 23 initial poses distant of 20 cm to the desired pose, the convergence rate of DDVS is 91% (21 successes, 2 failures in local minima). The convergence rate decreases as the sphere radius of initial poses increases: 60% for the radius 40 cm, 56% for the radius 60 cm. Interestingly, there is only one more failure from initial poses on the sphere of radius 60 cm than on the one of 40 cm: 10 failures for the former versus 9 failures for the latter.

Figure 6 shows the evolution of the cost [1, Eq. (15)] with respect to time for initial poses on the sphere of radius 20 cm. Figures 7 and 8 do the same for initial poses on spheres of radii 40 cm, resp. 60 cm.

As a summary, DDVS converges from 49 initial poses to the desired pose: a total convergence rate of $48/69 = 70\%$.

The penultimate column of Table I highlights for which initial poses DDVS succeeds or fails.

IV. PGM VS convergence domain

In order to show the contribution of DDVS over the state-of-the-art, PGM VS is run from the same 69 initial poses.

In detail, from the 23 initial poses distant of 20 cm to the desired pose, the convergence rate of PGM VS is 87% (20 successes, 3 failures in local minima). As DDVS, the convergence rate of PGM VS decreases as the sphere radius of initial poses increases, but more drastically: 35% for the radius 40 cm and even 0% for the radius 60 cm. In most of the cases, the camera diverges until reaching the limits of the robot workspace (Fig. 5d).

To sum up, PGM VS converges from 28 initial poses to the desired pose. The total convergence rate is $28/69 = 40\%$.

The right column of Table I highlights for which initial poses PGM VS succeeds or fails.

V. Comparison of convergence domains

To conclude, DDVS and PGM VS show almost identical convergence rates when the distance between the initial and the desired poses is 20 cm. However, when the latter distance increases to 40 cm and 60 cm, DDVS convergence rate is more than three times the one of PGM VS.

References

[1] G. Caron, "Defocus-based Direct Visual Servoing," *IEEE Robotics and Automation Letters*, 2021.

[2] N. Crombez, E. Mouaddib, G. Caron, and F. Chaumette, "Visual servoing with photometric gaussian mixtures as dense features," *IEEE T. on Rob.*, vol. 35, no. 1, pp. 49–63, 2019.

[3] C. Teulière and E. Marchand, "A Dense and Direct Approach to Visual Servoing Using Depth Maps," *IEEE Trans. on Robotics*, vol. 30, no. 5, pp. 1242–1249, 2014.

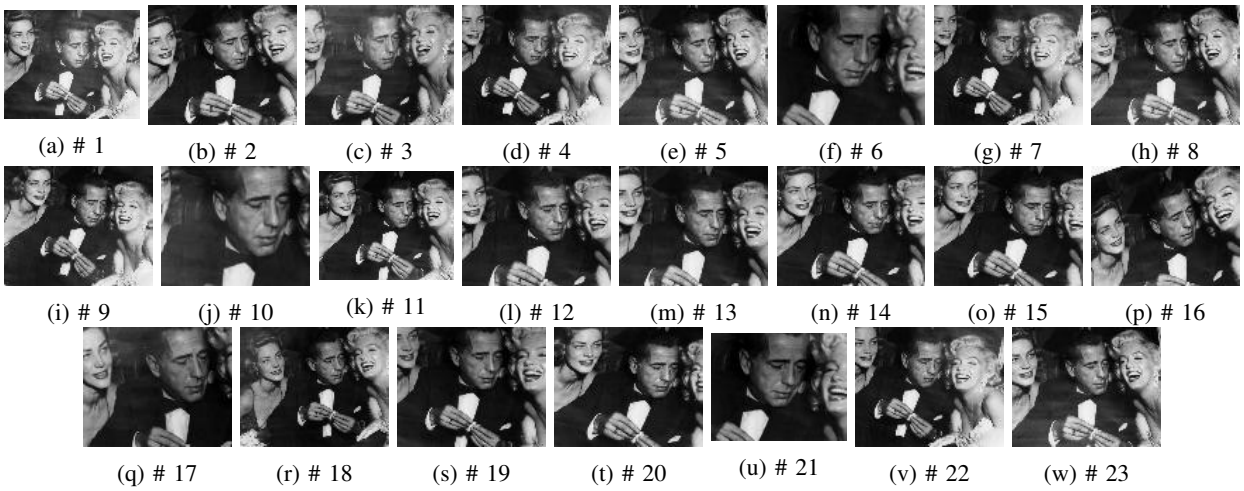


Fig. 1: Images captured at each of the 23 initial poses on the sphere of radius 20 cm. IDs are refering to those of Tab. I.

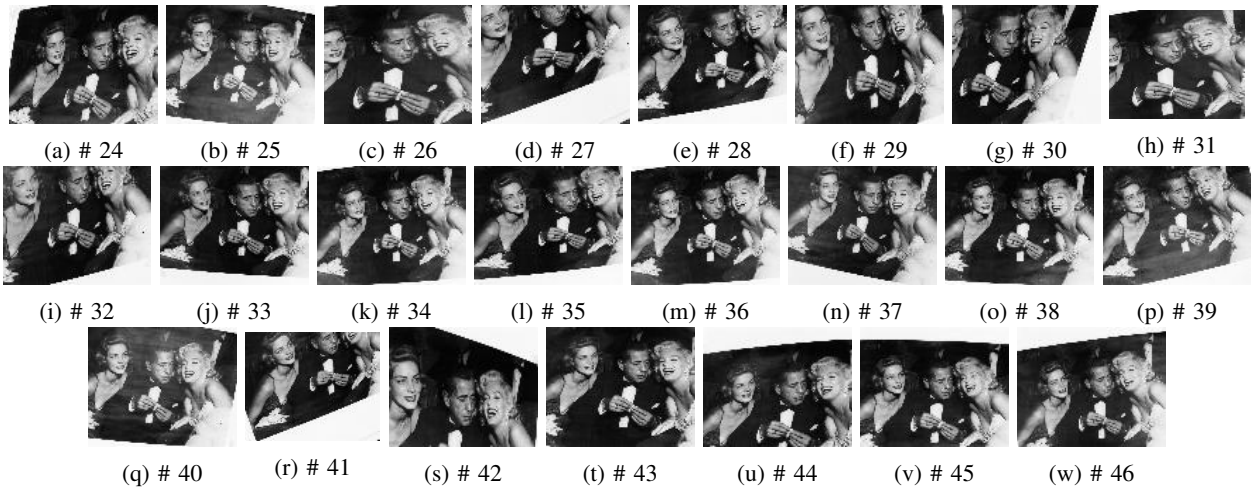


Fig. 2: Images captured at each of the 23 initial poses on the sphere of radius 40 cm. IDs are refering to those of Tab. I.

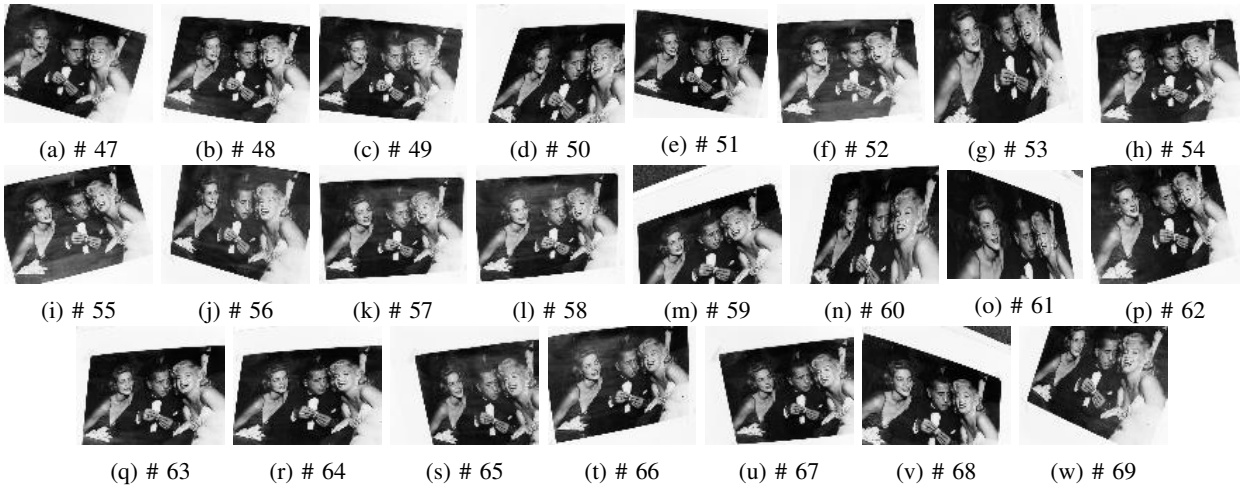


Fig. 3: Images captured at each of the 23 initial poses on the sphere of radius 60 cm. IDs are refering to those of Tab. I.

TABLE I: Initial camera poses relatively to the desired. Translations $t_X \in \mathbb{R}$, $t_Y \in \mathbb{R}$ and $t_Z \in \mathbb{R}$ are in cm. Rotations are represented as angle $\theta \in \mathbb{R}$ in degrees and unit axis $\mathbf{w} = [w_X, w_Y, w_Z] \in \mathbb{R}^3 : ||\mathbf{w}|| = 1$. The last two columns show the success (**o**) or failure (x) of DDVS (D) and PGM VS (G) from each initial pose.

#	t_X	t_Y	t_Z	θw_X	θw_Y	θw_Z	D	G
1	12.5	5.8	14.4	-5.2	11.0	-7.1	**o**	**o**
2	-0.5	-19.9	1.7	21.1	-0.6	2.5	**o**	**o**
3	17.3	10.0	-1.5	-11.7	19.6	-2.7	**o**	**o**
4	-8.4	-10.9	-14.5	17.1	-13.3	14.8	**o**	**o**
5	-10.9	1.0	16.8	-0.9	-9.3	14.0	**o**	**o**
6	1.0	4.5	-19.5	-8.4	1.9	-4.0	**o**	**o**
7	-3.0	-14.9	13.0	13.3	-2.7	-0.6	**o**	**o**
8	14.8	-2.4	-13.3	3.8	21.9	-13.0	**o**	x
9	-2.2	-7.8	18.3	6.5	-1.9	-14.7	**o**	**o**
10	-7.8	-1.0	-18.4	1.9	-13.9	-4.6	**o**	**o**
11	0.6	-16.6	11.2	15.2	0.5	-5.4	**o**	**o**
12	7.0	-18.1	-4.7	21.8	8.8	-4.5	**o**	**o**
13	10.4	-16.4	-4.8	19.9	13.0	-7.6	**o**	**o**
14	-0.6	-20.0	0.4	21.6	-0.6	-7.7	**o**	**o**
15	-1.0	-20.0	-0.0	21.8	-1.2	-11.5	**o**	**o**
16	-15.1	-13.1	-0.8	14.9	-17.0	13.0	**o**	**o**
17	-11.9	8.5	-13.7	-13.1	-18.1	3.5	**o**	**o**
18	-6.4	13.1	13.7	-11.6	-5.7	5.2	**o**	**o**
19	-6.4	17.5	-7.3	-22.3	-8.5	-6.0	**o**	x
20	-5.1	-19.1	-2.8	22.1	-6.2	-12.9	**o**	**o**
21	0.6	9.7	-17.5	-16.6	1.1	-13.4	x	x
22	2.6	18.1	8.0	-17.3	2.6	-11.9	x	**o**
23	16.1	-11.8	1.7	12.8	17.3	0.4	**o**	**o**
24	-34.7	-16.9	10.6	15.6	-29.8	-0.8	**o**	**o**
25	16.6	-6.9	35.7	4.6	11.0	-6.7	x	**o**
26	1.1	-39.5	-6.2	42.0	1.4	4.0	**o**	**o**
27	26.4	29.7	-4.2	-33.0	30.0	10.0	x	x
28	14.9	34.6	13.4	-28.7	13.2	7.6	x	x
29	40.0	0.7	-0.7	-0.8	39.0	6.8	x	x
30	33.1	22.4	-2.3	-25.1	34.7	-8.0	x	x
31	2.3	-39.8	-3.5	40.5	2.9	8.4	**o**	**o**
32	36.6	15.0	5.8	-15.0	33.3	4.2	x	x
33	-9.6	30.6	23.8	-22.5	-7.4	0.7	x	x
34	-7.3	-12.4	37.3	8.1	-4.8	5.4	**o**	x
35	5.2	36.3	16.1	-28.8	4.5	7.0	x	x
36	-7.3	-12.4	37.3	8.1	-4.8	5.4	**o**	x
37	-18.1	2.8	35.6	-1.9	-12.0	-7.7	**o**	x
38	1.2	-21.2	33.9	14.2	0.8	-1.0	**o**	**o**
39	20.7	-8.3	33.2	5.7	14.0	12.3	**o**	x
40	19.1	15.0	31.8	-10.4	13.1	-6.8	**o**	x
41	18.2	28.8	21.0	-22.1	14.4	15.0	**o**	x
42	38.9	-7.6	-5.5	9.7	41.1	-9.8	x	**o**
43	-3.7	-35.0	19.0	26.9	-3.1	-3.9	**o**	**o**
44	-15.9	-33.5	14.9	30.2	-13.8	1.2	**o**	x
45	7.4	-25.7	29.7	20.7	5.3	-0.4	**o**	**o**
46	-4.2	-3.5	39.6	5.1	-2.7	8.0	**o**	x
47	-11.8	24.3	53.6	-10.3	-6.5	-13.0	**o**	x
48	-22.8	-21.5	51.2	14.8	-12.7	-6.0	**o**	x
49	3.0	27.9	53.0	-12.3	1.7	-4.5	**o**	x
50	-52.0	-25.4	15.9	23.9	-38.3	-6.7	x	x
51	-19.1	39.4	41.0	-20.5	-11.9	-6.0	**o**	x
52	25.0	-10.4	53.6	8.6	13.6	4.0	**o**	x
53	60.0	1.0	-1.0	1.6	50.8	7.9	x	x
54	-10.9	-18.5	56.0	12.8	-5.9	7.0	**o**	x
55	31.1	-12.5	49.8	10.0	17.3	13.4	x	x
56	-48.8	7.3	34.2	-2.1	-30.1	-6.8	**o**	x
57	16.9	30.0	49.1	-14.0	9.7	2.6	**o**	x
58	28.3	27.2	45.3	-13.1	16.5	3.1	**o**	x
59	-26.8	-40.1	35.6	28.0	-17.4	13.3	**o**	x
60	-59.2	-6.3	-7.3	11.2	-54.2	-5.1	x	x
61	55.7	-11.4	-19.2	17.5	58.2	-1.7	x	x
62	54.7	-4.2	24.3	6.1	36.3	10.9	x	x
63	-25.2	-30.0	45.4	20.3	-14.8	1.9	**o**	x
64	-13.6	-43.2	39.4	28.6	-8.6	1.6	**o**	x
65	-48.1	19.5	30.1	-10.8	-31.0	10.5	x	x
66	36.5	29.1	37.7	-15.5	22.6	6.4	x	x
67	-9.4	34.4	48.2	-16.4	-5.5	9.4	x	x
68	44.1	-30.9	26.5	22.0	30.0	-11.9	**o**	x
69	-56.6	-4.2	19.5	6.3	-39.2	-14.2	x	x

Fig. 4: The desired image acquired at the desired pose (center of spheres of initial poses). Same as [1, Fig. 8a].

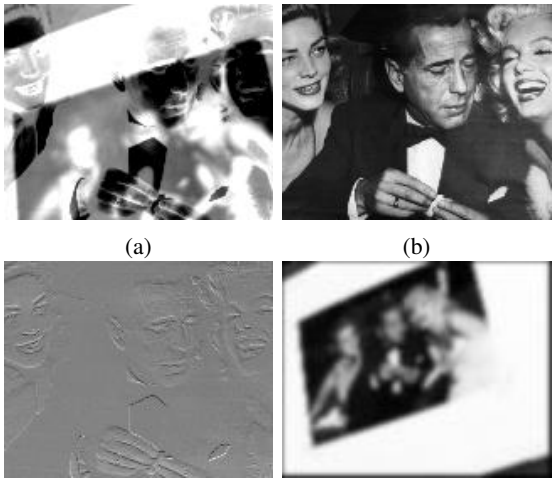


Fig. 5: Examples of result images for the initial pose # 59 (see Tab I). (a) Difference between the initial image (Fig. 3m) and the desired one (Fig. 4). (b) Last image of DDVS. (c) Difference between the last image of DDVS and the desired one. (d) Last PGM of PGM VS (the camera is on the border of the robot workspace).

[4] Q. Bateux and E. Marchand, "Histograms-Based Visual Servoing," *IEEE Robotics and Automation Letters*, vol. 2, no. 1, pp. 80–87, 2017.

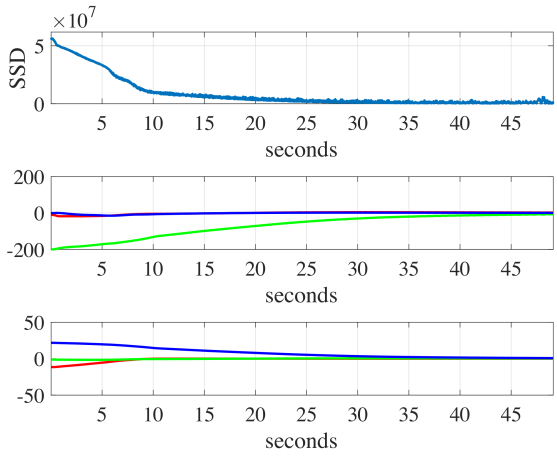
[5] M. Bakthavatchalam, O. Tahri, and F. Chaumette, "A Direct Dense Visual Servoing Approach using Photometric Moments," *IEEE T. on Rob.*, vol. 34, no. 5, pp. 1226–1239, 2018.

[6] E. Marchand, "Direct visual servoing in the frequency domain," *IEEE Rob. and Autom. Letters*, vol. 5, no. 2, pp. 620–627, 2020.

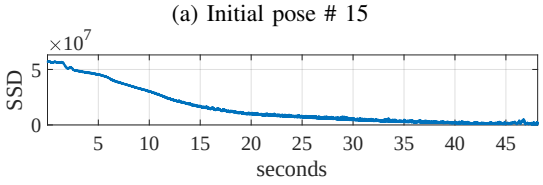
[7] G. Silveira, L. Mirisola, and P. Morin, "Decoupled Intensity-Based Nonmetric Visual Servo Control," *IEEE T. on Control Syst. Tech.*, vol. 28, no. 2, pp. 566–573, 2020.

[8] E. Marchand, F. Spindler, and F. Chaumette, "Visp for visual servoing: a generic software platform with a wide class of robot control skills," *IEEE Rob. and Autom. Magazine*, vol. 12, no. 4, pp. 40–52, December 2005.

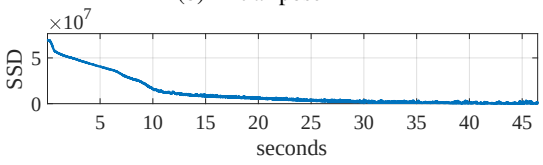
[9] M. ONeill, "PCG: A family of simple fast space-efficient statistically good algorithms for random number generation," Harvey Mudd College, Claremont, CA, Tech. Rep. HMC-CS-2014-0905, September 2014.



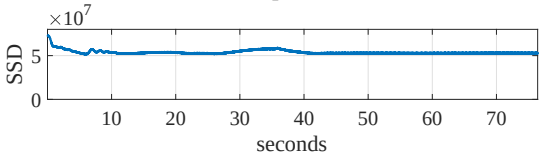
(a) Initial pose # 15



(b) Initial pose # 4

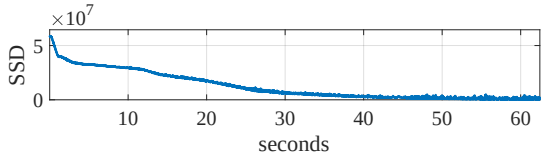


(c) Initial pose # 20

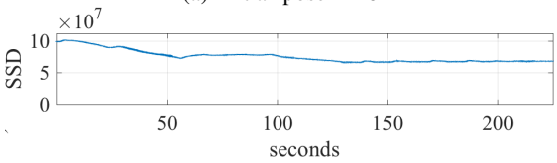


(d) Initial pose # 21 (divergence)

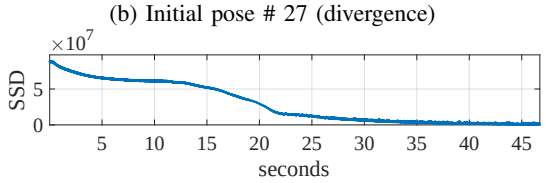
Fig. 6: Evolution of the Sum-of-Squared-Differences cost (SSD) between images brightness during the Defocus-based DVS from several initial poses on the sphere of radius 20 cm. (a) also shows the evolution of translational DoF (red: t_X, green : t_Y, blue: t_Z, unit: mm) on its second row and rotational DoF (red: θw_Z, green : θw_Y, blue: θw_X, unit: degrees) to highlight there smoothness.



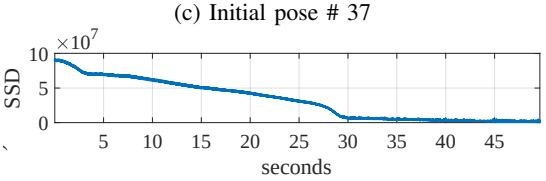
(a) Initial pose # 26



(b) Initial pose # 27 (divergence)

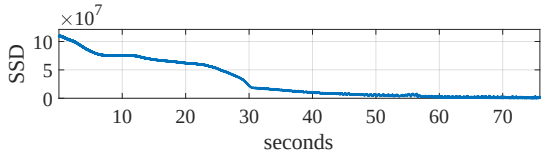


(c) Initial pose # 37

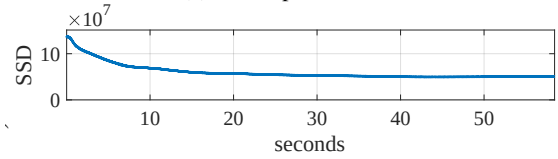


(d) Initial pose # 39

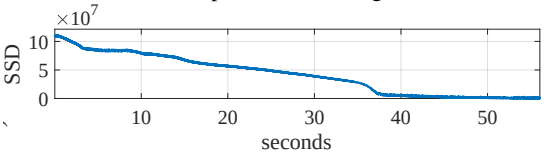
Fig. 7: Evolution of the Sum-of-Squared-Differences cost (SSD) between images brightness during the Defocus-based DVS from several initial poses on the sphere of radius 40 cm.



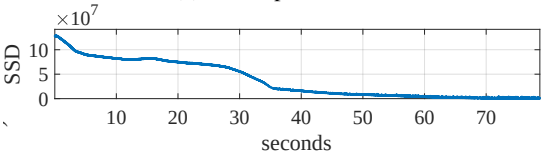
(a) Initial pose # 48



(b) Initial pose # 50 (divergence)



(c) Initial pose # 57



(d) Initial pose # 63

Fig. 8: Evolution of the Sum-of-Squared-Differences cost (SSD) between images brightness during the Defocus-based DVS from several initial poses on the sphere of radius 60 cm.