Supplementary material: Particle dynamics and multi-channel feature dictionaries for robust visual tracking

Srikrishna Karanam
karans3@rpi.edu
Yang Li
yangli625@gmail.com
Richard J. Radke
rjradke@ecse.rpi.edu

Department of Electrical, Computer, and Systems Engineering
Rensselaer Polytechnic Institute
110 8th St.
Troy, NY USA

1 Additional results

The mean CLE and success rate for each individual test sequence are shown in Tables 1 and 2 respectively. The success plots for the following attributes: background clutter (BC), motion blur (MB), fast motion (FM), occlusion (OCC), non-rigid object deformation (DEF), out-of-view (OV), and low resolution are shown in Figure 1.

2 Adaptive candidate filtering

We first derive the expression representing the number of particles to be chosen in each frame:

$$n = \frac{1}{2\nu}\chi^2_{k-1,1-\delta} \approx \frac{k-1}{2\nu}\left(1 - \frac{2}{9(k-1)} + \sqrt{\frac{2}{9(k-1)}}z_{1-\delta}\right)^3 \tag{1}$$

Consider two probability distributions p_1 and p_2. The Kullback-Leibler distance [2] K between p_1 and p_2 is defined as

$$K(p_1, p_2) = \sum_x p_1(x)\log\left(\frac{p_1(x)}{p_2(x)}\right) \tag{2}$$

The basic idea of KLD-sampling [3] is to find the number of particles in each iteration such that the error between the true posterior probability density and the probability density approximated by the particle filter is less than ν with probability $(1-\delta)$. At any particular iteration, suppose we draw n particles from a discrete probability distribution that has k disparate bins. Defining the vector $\mathbf{N} = [N_1, N_2, \ldots, N_k]$ as the number of particles drawn from each bin, we can see that $\mathbf{N}$ follows a multinomial distribution $f_k(n, \mathbf{p})$, where $\mathbf{p} = [p_1, p_2, \ldots, p_k]$ represents the probability of each of the k bins. We can use the maximum

Table 1: Mean center location error (in pixels) for each of the 25 test sequences. **Red** - Best, *Blue* - Second best.

Sequence	Ours	L1	MTT	ONDL	SCM	LSH	ASLA	SPT	LOT	MIL	IVT
Basketball	**6.17**	128	94.3	8.80	55.4	7.65	*6.40*	13.7	6.57	97.4	86.6
Boy	4.44	4.05	15.9	*2.74*	51.1	6.37	**2.18**	4.93	66	12.8	91.8
Car4	2.7	85.0	22.8	2.26	4.05	54.8	**1.70**	98.1	167	50.7	*2.04*
CarScale	*5.27*	82.7	74.2	15.7	28.8	10.2	20.8	**3.9**	91.8	31.6	10.3
Coke	**8.7**	117	24.8	67.2	49.1	73.9	60.9	*11.8*	62.4	43.6	83.1
Crossing	**1.30**	2.8	56.3	1.85	*1.31*	50.3	1.67	39.7	36.7	3.04	2.6
David	**6.8**	54.4	10	23.7	10	14	*6.82*	27.1	38.5	17.7	9.21
David2	1.61	15.3	**1.27**	3.96	3.81	2.69	*1.36*	46.6	4.1	10.9	1.43
Deer	**5.27**	163	8.97	7.87	12.2	7.69	**4.96**	36.3	97.5	101	182
Dudek	**9.29**	33.5	14.7	10.1	10.7	12.5	14.9	70.2	85.1	17.7	*9.49*
FaceOcc2	*6.7*	13.6	8.9	**5.9**	15.5	11.5	19	21.4	15	13.6	7.1
FleetFace	**15.5**	26.3	69	*19.4*	27.8	28.6	31.7	234	33.7	63.1	62.5
Football1	*5.1*	12.8	13.1	8.4	20	**5.09**	11.6	48.3	6.85	5.62	24.3
Girl	**3.5**	*5.1*	9	37.9	64.6	37.1	6.3	10.6	21.4	13.8	22.6
MountainBike	**6**	210	7.3	*6.58*	10.4	7.8	8.8	11.8	24.9	73	7.4
Shaking	*7.64*	125	97.2	**7.21**	10.8	8.04	22.7	130	82.6	24	85.3
Singer1	**2.76**	3.36	35.1	3.33	3.28	14.5	*2.87*	80.5	140	16.5	11.5
Singer2	**7.04**	184	210	179	113	*8.71*	175	225	76.9	22.5	175
Skating1	**6.02**	132	298	*7.12*	9.21	68.2	48.6	188	88.5	139	146
Soccer	**19.1**	129	84.3	89.4	77.8	101	119	53.8	*42.2*	77.8	145
Sylvester	12.7	49.8	*7.37*	7.54	8.08	**6.45**	15.3	33.6	11.4	15.4	34.3
Trellis	**2.98**	84.9	59.3	18.3	*5.61*	32.7	7.81	11.69	47.6	71.5	119
Walking	2.26	3.5	2.76	3.45	2.44	12.8	*2.02*	37.3	2.42	3.42	**1.71**
Walking2	*2.37*	2.67	3.36	36.1	**1.85**	23.4	37.7	25.7	64.7	60.4	2.76
Woman	**3.38**	356	105	5.7	123	7.52	*3.71*	7.44	114	102	142

Table 2: Mean success rate for each of the 25 test sequences. **Red** - Best, *Blue* - Second best.

Sequence	Ours	L1	MTT	ONDL	SCM	LSH	ASLA	SPT	LOT	MIL	IVT
Basketball	**96.5**	10.8	25.2	90.2	31.2	94.3	*95.1*	80.5	89.2	30.6	10.3
Boy	94.8	96.8	44.5	97.8	43.8	89.2	*99.5*	**100**	64.9	38.5	32.5
Car4	**100**	15.6	31.4	**100**	*97.2*	27.2	**100**	19.6	4.8	27.6	*100*
CarScale	**100**	68	57.7	74.3	68	46.9	73.0	*98.8*	48.6	46.9	73.9
Coke	**96.8**	10.4	69.6	23.2	40.8	6.8	15.2	*87.2*	10.4	12.4	15.2
Crossing	**100**	35.8	22.5	**100**	**100**	12.5	100	35.8	30.8	*98.3*	24.2
David	**90**	24.7	*85.5*	36	79	47.5	80.5	9	2.75	20	68
David2	**100**	72.8	**100**	75	91.3	100	*94.6*	27.6	76.9	32.4	92.4
Deer	**100**	5.63	*95.7*	**100**	92.9	94.4	**100**	52.1	2.82	12.7	2.82
Dudek	**100**	71.8	92.7	96.9	97.5	*97.6*	89.8	56.8	61.8	85.7	96.9
FaceOcc2	*98.3*	72.7	90.6	**99.7**	71.5	97.2	43.2	65.5	35	93.6	91.6
FleetFace	**94.3**	79.6	54.7	*83.6*	70.6	71.4	59.5	0.6	57.8	53.7	46.5
Football1	*75.7*	28.4	59.5	54	41.9	**85.1**	44.6	17.6	41.9	78.4	49.5
Girl	**91.4**	62.6	62.4	24	34.2	15.4	*85.6*	53.8	58.6	29.4	18.4
MountainBike	**100**	28.5	95.2	95.2	*96.9*	**100**	89.9	36.8	68.8	57.5	**100**
Shaking	**100**	0.5	1.1	94.8	90.1	*95.9*	32.6	8.49	7.67	22.7	1.1
Singer1	**100**	**100**	35.6	*99.7*	**100**	27.6	**100**	23.4	24.8	27.6	44.2
Singer2	*97.5*	3.55	3.55	3.55	16.4	**100**	3.55	3.28	15.8	47.5	3.8
Skating1	**99**	9.25	13	41.7	35.2	9.25	*51.7*	19	24	10.2	9.5
Soccer	**33**	16.3	18.1	17.9	23.5	8.9	12.5	*26*	21.7	15.6	17.3
Sylvester	82.4	29.5	82.2	85.2	*88.5*	**96.3**	74.8	24.5	67.7	54.6	67.4
Trellis	**100**	21.3	23.7	79.6	*96.5*	44.1	85	74.5	33	24.4	30.9
Walking	**99.7**	96.6	*99*	*99*	96.1	29.8	**99.7**	43	96.8	54.1	**99.7**
Walking2	99.2	99.6	99.2	40.6	**100**	38.8	39.8	29.6	39	38	*99.8*
Woman	**99.7**	23.5	29.7	90.5	30	*97*	96.2	95.5	15.5	28	27.7

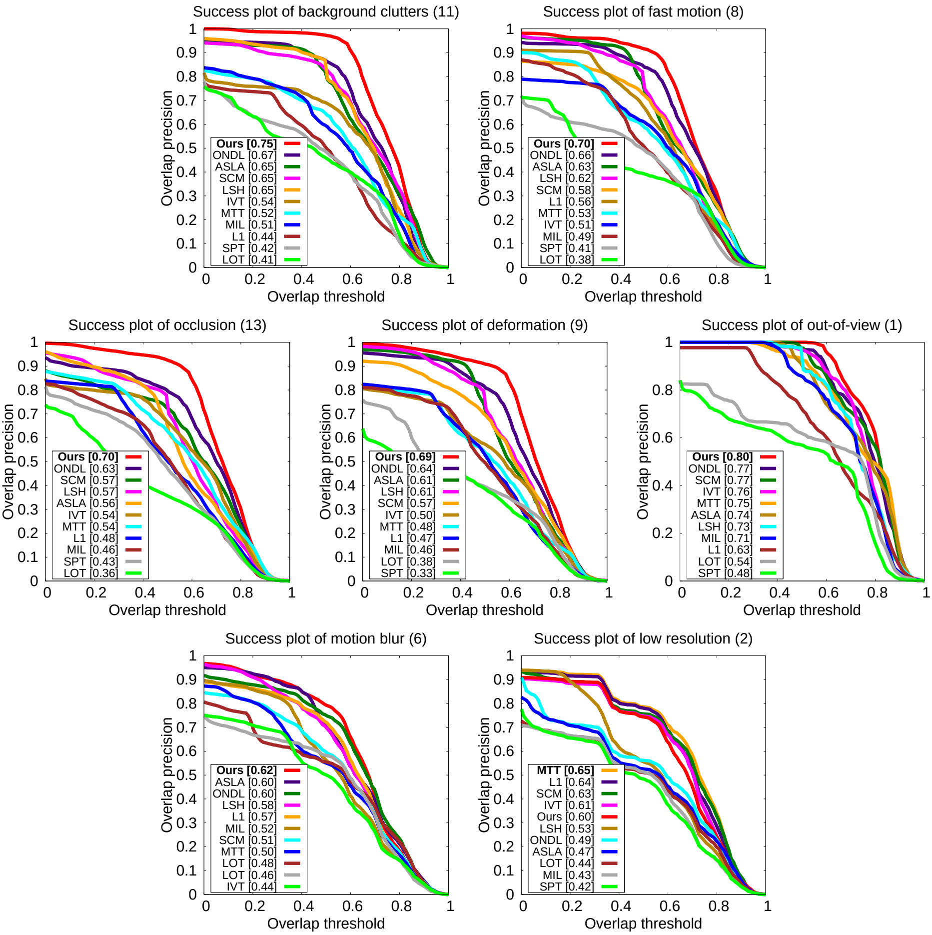

Figure 1: Success plots for background clutter, fast motion, occlusion, non-rigid object deformation, out-of-view, motion blur, fast motion, and low-resolution attributes.

likelihood estimation procedure to obtain $\hat{\mathbf{p}}$ as

$$\hat{\mathbf{p}} = \frac{\mathbf{N}}{n} \tag{3}$$

The likelihood ratio λ_n statistic for $\mathbf{p}$ is given by

$$\log \lambda_n = \sum_{j=1}^{k} N_j \log \frac{\hat{p}_j}{p_j} \tag{4}$$

Since $N_j = n\hat{p}_j$, this equation becomes

$$\log \lambda_n = n \sum_{j=1}^{k} \hat{p}_j \log \frac{\hat{p}_j}{p_j} = n\mathrm{K}(\hat{\mathbf{p}}, \mathbf{p}) \tag{5}$$

Noting that $2\log \lambda_n$ converges in distribution to a chi-square distribution as $n \to \infty$ [4], consider the probability $P(\mathrm{K}(\hat{\mathbf{p}}, \mathbf{p}) \leq \nu)$:

$$\begin{aligned} P(\mathrm{K}(\hat{\mathbf{p}}, \mathbf{p}) \leq \nu) &= P(2n\mathrm{K}(\hat{\mathbf{p}}, \mathbf{p}) \leq 2n\nu) \\ &= P(2\log \lambda_n \leq 2n\nu) \\ &= P\left(\chi^2_{k-1} \leq 2n\nu\right) \end{aligned} \tag{6}$$

Using the fact that $P\left(\chi^2_{k-1} \leq \chi^2_{k-1,1-\delta}\right) = 1 - \delta$, if we choose n according to the following expression:

$$2n\nu = \chi^2_{k-1,1-\delta} \tag{7}$$

we get

$$P(\mathrm{K}(\hat{\mathbf{p}}, \mathbf{p}) \leq \nu) = 1 - \delta \tag{8}$$

which is exactly what we wished to achieve, hence completing the proof. We see that equation 1 follows from Equation 7.

3 Optimization problem

In each feature channel, we solve the following optimization problem:

$$\begin{aligned} \min_{\mathbf{x}^j, \varepsilon^j} \quad & \|\mathbf{x}^j\|_1 + \|\varepsilon^j\|_1 \\ \text{s.t.} \quad & \mathbf{y}^j = \mathbf{A}^j \mathbf{x}^j + \varepsilon^j \end{aligned} \tag{9}$$

This problem is of the general form

$$\begin{aligned} \min_{\mathbf{x}, \varepsilon} \quad & f_1(\mathbf{x}, \varepsilon) \\ \text{s.t.} \quad & f_2(\mathbf{x}, \varepsilon) = \mathbf{0} \end{aligned} \tag{10}$$

where $f_2(\mathbf{x}, \varepsilon) = \mathbf{y} - \mathbf{A}\mathbf{x} - \varepsilon$. Both $f_2(\mathbf{x}, \varepsilon)$, and $f_1(\mathbf{x}, \varepsilon)$ are continuous and convex functions in $(\mathbf{x}, \varepsilon)$, and hence the problem

$$\begin{aligned} \min_{\mathbf{x}, \varepsilon} \quad & f_1(\mathbf{x}, \varepsilon) + \tfrac{\zeta}{2} \|f_2(\mathbf{x}, \varepsilon)\|_2^2 \\ \text{s.t.} \quad & f_2(\mathbf{x}, \varepsilon) = \mathbf{0} \end{aligned} \tag{11}$$

has the same optimal value pair $(\mathbf{x}^*, \varepsilon^*)$ as the problem defined in Equation 10. We now eliminate the equality constraints in this problem by introuducing the Lagrange multipliers. The *augmented* Lagrangian for this problem is

$$\mathcal{L}_\zeta(\mathbf{x}, \varepsilon, \rho) = f_1(\mathbf{x}, \varepsilon) + \frac{\zeta}{2}\|f_2(\mathbf{x}, \varepsilon)\|_2^2 + \rho^T f_2(\mathbf{x}, \varepsilon) \tag{12}$$

The minimization problem of Equation 11 is equivalent to minimizing the augmented Lagrangian of Equation 12. Therefore, we now have

$$(\mathbf{x}^*, \varepsilon^*) = \arg\min_{\mathbf{x},\varepsilon} \mathcal{L}_\zeta(\mathbf{x}, \varepsilon, \rho) \tag{13}$$

The minimization problem of Equation 13 can be solved using the framework of alternating directions algorithms [6]. Specifically, in each iteration, we compute $\mathbf{x}$ and ε separately, and then update ρ. Formally, the optimal solution pair $(\mathbf{x}^*, \varepsilon^*)$ is computed as

$$\mathbf{x}_{i+1} = \arg\min_{\mathbf{x}} \mathcal{L}_\zeta(\mathbf{x}, \varepsilon_i, \rho_i) \tag{14}$$

$$\varepsilon_{i+1} = \arg\min_{\varepsilon} \mathcal{L}_\zeta(\mathbf{x}_{i+1}, \varepsilon, \rho_i) \tag{15}$$

$$\rho_{i+1} = \rho_i + \zeta(f_2(\mathbf{x}_{i+1}, \varepsilon_{i+1})) \tag{16}$$

The sub-problem defined by

$$\varepsilon_{i+1} = \arg\min_{\varepsilon} \mathcal{L}_\zeta(\mathbf{x}_{i+1}, \varepsilon, \rho_i) \tag{17}$$

has a closed form solution, which we derive next. Consider the definition

$$\mathcal{L}_\zeta(\mathbf{x}_{i+1}, \varepsilon, \rho_i) = \|\mathbf{x}_{i+1}\|_1 + \|\varepsilon\|_1 + \frac{\zeta}{2}\|f_2(\mathbf{x}_{i+1}, \varepsilon)\|_2^2 + \rho_i^T(f_2(\mathbf{x}_{i+1}, \varepsilon)) \tag{18}$$

Defining $\varepsilon_\mathbf{d} = \mathbf{y} - \mathbf{A}\mathbf{x}_{i+1}$, minimizing $\mathcal{L}_\zeta(\mathbf{x}_{i+1}, \varepsilon, \rho_i)$ is equivalent to

$$\begin{aligned}
\varepsilon^* &= \arg\min_{\varepsilon} \{\|\varepsilon\|_1 + \frac{\zeta}{2}\|f_2(\mathbf{x}_{i+1}, \varepsilon)\|_2^2 + \rho_i^T(f_2(\mathbf{x}_{i+1}, \varepsilon))\} \\
&= \arg\min_{\varepsilon} \{\|\varepsilon\|_1 + \rho_i^T(\varepsilon_d - \varepsilon) + \frac{\zeta}{2}(\varepsilon_d - \varepsilon)^T(\varepsilon_d - \varepsilon)\} \\
&= \arg\min_{\varepsilon} \left\{\|\varepsilon\|_1 + \frac{\zeta}{2}\left\|\varepsilon - \left(\varepsilon_d + \frac{\rho_i}{\zeta}\right)\right\|_2^2\right\} \\
&= \mathcal{T}_{\frac{1}{\zeta}}\left(\varepsilon_d + \frac{\rho_i}{\zeta}\right)
\end{aligned} \tag{19}$$

where $\mathcal{T}_\alpha(\mathbf{t})_i = \text{sgn}(t_i)\max\{|t_i| - \alpha, 0\}$, $i = 1, 2, \dots, n$. Thus, the update step for ε_{i+1} has an analytic solution given by

$$\varepsilon_{i+1} = \mathcal{T}_{\frac{1}{\zeta}}\left(\varepsilon_\mathbf{d} + \frac{\rho_i}{\zeta}\right) \tag{20}$$

However, the sub-problem defined by

$$\mathbf{x}_{i+1} = \arg\min_{\mathbf{x}} \mathcal{L}_\zeta(\mathbf{x}, \varepsilon_i, \rho_i) \tag{21}$$

does not have an analytic solution, and we hence must resort to iterative schemes. To solve this problem, we use the Fast Iterative Shrinkage Thresholding Algorithm (FISTA) [1]. We first show that this optimization problem is basically the classic lasso [5] problem. Defining $\mathbf{b}^{'} = \mathbf{y} - \varepsilon_i$, and $\mathbf{b}^{''} = \mathbf{b}^{'} + \frac{\rho_i}{\zeta}$, we have

$$\begin{aligned} \mathbf{x}_{i+1} &= \arg\min_{\mathbf{x}} \{\|\mathbf{x}\|_1 + \rho_i{}^T(\mathbf{b}^{'} - \mathbf{A}\mathbf{x}) + \frac{\zeta}{2}(\mathbf{b}^{'} - \mathbf{A}\mathbf{x})^T(\mathbf{b}^{'} - \mathbf{A}\mathbf{x})\} \\ &= \arg\min_{\mathbf{x}} \left\{\|\mathbf{x}\|_1 + \frac{\zeta}{2}\left\|\mathbf{A}\mathbf{x} - \mathbf{b}^{'} - \frac{\rho_i}{\zeta}\right\|_2^2\right\} \\ &= \arg\min_{\mathbf{x}} \left\{\|\mathbf{x}\|_1 + \frac{\zeta}{2}\|\mathbf{A}\mathbf{x} - \mathbf{b}^{''}\|_2^2\right\} \end{aligned} \tag{22}$$

Thus, we see that the problem of Equation 21 reduces to the lasso framework, which can be efficiently solved using FISTA.

References

[1] Amir Beck and Marc Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIIMS*, 2(1):183–202, March 2009.

[2] Thomas M Cover and Joy A Thomas. *Elements of information theory*. John Wiley & Sons, 2012.

[3] Dieter Fox. Kld-sampling: Adaptive particle filters. In *NIPS*, pages 713–720, 2001.

[4] John A Rice. *Mathematical Statistics and Data Analysis*. Duxbury Advanced Series, Belmont, CA, 2006.

[5] Robert Tibshirani. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, pages 267–288, 1996.

[6] Junfeng Yang and Yin Zhang. Alternating direction algorithms for l_1-problems in compressive sensing. 33(1):250–278, February 2011.