A Sparse Gaussian Approach to Region-Based 6DoF Object Tracking – Supplementary

Manuel Stoiber[1,2][0000−0002−0762−9288], Martin Pfanne[1][0000−0003−2076−4772], Klaus H. Strobl[1][0000−0001−8123−0606], Rudolph Triebel[1,2][0000−0002−7975−036X], and Alin Albu-Schäffer[1,2][0000−0001−5343−9074]

[1] German Aerospace Center (DLR), 82234 Wessling, Germany
manuel.stoiber@dlr.de
[2] Technical University of Munich (TUM), 80333 Munich, Germany

A Mathematical Proof

In the following, we provide a detailed version of the mathematical proof discussed in Section 4.1. The goal is to find smoothed step functions h_f and h_b that ensure that the likelihood follows a Gaussian distribution. For this, we first state our assumptions and simplify the likelihood function. The first-order derivative of the log-likelihood is then calculated to eliminate constant scaling terms and find a concise expression. Finally, we enforce that the first-order derivative of the logarithm of the likelihood has to be equal to that of a normal distribution. As a result, we obtain the required smoothed step functions and the proof that with those functions the likelihood follows a Gaussian distribution.

For the proof, we start from the likelihood function that was derived in Section 4.1 and that was defined as follows

$$p(\boldsymbol{\mathcal{D}}_i|\boldsymbol{\theta}) \propto \prod_{r\in\mathcal{R}_i} \Big(h_\mathrm{f}(r - \Delta c_i^+)p_{\mathrm{f}i}(r) + h_\mathrm{b}(r - \Delta c_i^+)p_{\mathrm{b}i}(r)\Big), \tag{1}$$

with $\boldsymbol{\mathcal{D}}_i$ the data specific to a single correspondence line, $\mathcal{R}_i$ a set of distances r from the line center to pixel centers that ensures that every pixel along the line appears exactly once, h_f and h_b the smoothed step functions for foreground and background, $p_{\mathrm{f}i}$ and $p_{\mathrm{b}i}$ the pixel-wise posteriors for foreground and background, and Δc_i^+ the projected difference from the correspondence line center $\boldsymbol{c}_i$ to the variated model point ${}_\mathrm{C}\boldsymbol{X}_i^+$.

For pixel-wise posteriors, perfect segmentation and a contour at the correspondence line center are assumed. This results in the following step functions

$$p_{\mathrm{f}i}(r) = \begin{cases} 1 & \text{if } r \leq 0 \\ 0 & \text{else} \end{cases}, \quad p_{\mathrm{b}i}(r) = \begin{cases} 0 & \text{if } r \leq 0 \\ 1 & \text{else} \end{cases}. \tag{2}$$

Also, we restrict the smoothed step functions h_f and h_b to sum to one and to be symmetric. Consequently, the following expressions can be defined

$$h_\mathrm{f}(x) = 0.5 - f(x), \quad h_\mathrm{b}(x) = 0.5 + f(x), \tag{3}$$

where $f(x)$ is an odd function that lies within the interval $[-0.5, 0.5]$ and that fulfills $\lim_{x\to\infty} f(x) = 0.5$ and $\lim_{x\to-\infty} f(x) = -0.5$. Finally, infinitesimally small pixels are assumed to write the likelihood from Eq. (1) in continuous form

$$p(\boldsymbol{\mathcal{D}}_i|\boldsymbol{\theta}) \propto \prod_{r=-\infty}^{\infty} \Big(h_\mathrm{f}(r - \mathit{\Delta} c_i^+)p_{\mathrm{f}i}(r) + h_\mathrm{b}(r - \mathit{\Delta} c_i^+)p_{\mathrm{b}i}(r)\Big)^{dr}. \tag{4}$$

Having stated all assumptions, we start by simplifying Eq. (4). For this, the product integral is first converted to the classical Riemann integral

$$p(\boldsymbol{\mathcal{D}}_i|\boldsymbol{\theta}) \propto \exp\bigg(\int_{r=-\infty}^{\infty} \ln\Big(h_\mathrm{f}(r - \mathit{\Delta} c_i^+)p_{\mathrm{f}i}(r) + h_\mathrm{b}(r - \mathit{\Delta} c_i^+)p_{\mathrm{b}i}(r)\Big)dr\bigg). \tag{5}$$

The integral is then split at $r = 0$ and the definitions from Eq. (2) are used

$$p(\boldsymbol{\mathcal{D}}_i|\boldsymbol{\theta}) \propto \exp\bigg(\int_{r=-\infty}^{0} \ln\big(h_\mathrm{f}(r - \mathit{\Delta} c_i^+)\big)dr + \int_{r=0}^{\infty} \ln\big(h_\mathrm{b}(r - \mathit{\Delta} c_i^+)\big)dr\bigg). \tag{6}$$

Finally, $x = r - \mathit{\Delta} c_i^+$ is substituted to write the following simplified expression

$$p(\boldsymbol{\mathcal{D}}_i|\boldsymbol{\theta}) \propto \exp\bigg(\int_{x=-\infty}^{-\mathit{\Delta} c_i^+} \ln\big(h_\mathrm{f}(x)\big)dx + \int_{x=-\mathit{\Delta} c_i^+}^{\infty} \ln\big(h_\mathrm{b}(x)\big)dx\bigg). \tag{7}$$

To eliminate both constant scaling factors and the integral, we first apply the logarithm and then use Leibniz's rule for differentiation under the integral to calculate the first-order derivative with respect to $\mathit{\Delta} c_i^+$

$$\frac{\partial \ln\big(p(\boldsymbol{\mathcal{D}}_i|\boldsymbol{\theta})\big)}{\partial \mathit{\Delta} c_i^+} = -\ln\big(h_\mathrm{f}(-\mathit{\Delta} c_i^+)\big) + \ln\big(h_\mathrm{b}(-\mathit{\Delta} c_i^+)\big). \tag{8}$$

Note that definitions for the smoothed step functions were used to ensure that $\lim_{x\to-\infty} \ln(h_\mathrm{f}(x)) = 0$ and $\lim_{x\to\infty} \ln(h_\mathrm{b}(x)) = 0$. Introducing Eq. (3), one gets the following equation that only depends on a single unknown function

$$\begin{aligned} \frac{\partial \ln\big(p(\boldsymbol{\mathcal{D}}_i|\boldsymbol{\theta})\big)}{\partial \mathit{\Delta} c_i^+} &= -\ln\big(0.5 - f(-\mathit{\Delta} c_i^+)\big) + \ln\big(0.5 + f(-\mathit{\Delta} c_i^+)\big) && (9)\\ &= -\ln\big(1 - 2f(-\mathit{\Delta} c_i^+)\big) + \ln\big(1 + 2f(-\mathit{\Delta} c_i^+)\big). && (10) \end{aligned}$$

The definition of the inverse hyperbolic tangent

$$2\tanh^{-1}(z) = -\ln(1 - z) + \ln(1 + z), \tag{11}$$

is then used to obtain a concise expression

$$\frac{\partial \ln\big(p(\boldsymbol{\mathcal{D}}_i|\boldsymbol{\theta})\big)}{\partial \mathit{\Delta} c_i^+} = 2\tanh^{-1}\big(2f(-\mathit{\Delta} c_i^+)\big). \tag{12}$$

In the following, we want to enforce that our likelihood function follows a normal distribution. We thus start with the definition of a normal distribution

$$\mathcal{N}(\Delta c_i^+|0, s_\mathrm{h}) \propto \exp\Big(-\frac{1}{2s_\mathrm{h}}(\Delta c_i^+)^2\Big), \tag{13}$$

where s_h describes the variance of the normal distribution. To achieve our goal, the first-order derivative of the log-likelihood in Eq. (12) is set equal to the first-order derivative of the logarithm of the normal distribution

$$\frac{\partial \ln\big(p(\boldsymbol{\mathcal{D}}_i|\boldsymbol{\theta})\big)}{\partial \Delta c_i^+} = \frac{\partial \ln\big(\mathcal{N}(\Delta c_i^+|0, s_\mathrm{h})\big)}{\partial \Delta c_i^+} \tag{14}$$

$$2\tanh^{-1}\big(2f(-\Delta c_i^+)\big) = -\frac{\Delta c_i^+}{s_\mathrm{h}}. \tag{15}$$

Solving for f, one obtains the following expression

$$f(x) = \frac{1}{2}\tanh\Big(\frac{x}{2s_\mathrm{h}}\Big). \tag{16}$$

Introducing Eq. (16) into the original definitions from Eq. (3), the final expressions for our smoothed step functions can be written as follows

$$h_\mathrm{f}(x) = \frac{1}{2} - \frac{1}{2}\tanh\Big(\frac{x}{2s_\mathrm{h}}\Big), \quad h_\mathrm{b}(x) = \frac{1}{2} + \frac{1}{2}\tanh\Big(\frac{x}{2s_\mathrm{h}}\Big). \tag{17}$$

Notice that the variance of the designed likelihood function s_h turned into a slope parameter. The used equality constrained between the first-order derivative of the log-likelihood and the first-order derivative of the logarithm of the normal distribution ensures that the original functions can only differ by a constant scaling factor. We are thus able to write

$$p(\boldsymbol{\mathcal{D}}_i|\boldsymbol{\theta}) \propto \mathcal{N}(\Delta c_i^+|0, s_\mathrm{h}). \tag{18}$$

This shows that, for the derived smoothed step functions with the slope parameter s_h, one obtains a likelihood function that is proportional to a normal distribution with the variance s_h.

B Ablation Study

In the following, we present an ablation study that assesses the importance of individual components of our approach. The conducted experiments form the basis for our remarks on the algorithm's performance in Section 6.3. The evaluation is conducted on the *noise* sequence of the *RBOT* dataset, which features dynamic lighting, Gaussian noise, and simulated motion blur. The experiments, as well as the calculation of the success rate, are performed in the exact same manner as described in Section 6.1. For the final result, the average success rate over all 18 objects of the *RBOT* dataset is calculated.

In our study, we compare the original tracker to three experiments that feature different tracker configurations. In *Experiment 1*, we set $\lambda_\mathrm{r} = 0$ and $\lambda_\mathrm{t} = 0$ to evaluate the tracker without Tikhonov regularization. *Experiment 2* features linear smoothed step functions h_f and h_b, where h_f linearly decreases from $h_\mathrm{f}(-4.5) = 1$ to $h_\mathrm{f}(4.5) = 0$ and h_b linearly increases from $h_\mathrm{b}(-4.5) = 0$ to $h_\mathrm{b}(4.5) = 1$. While an infinite number of functions would be possible, the experiment qualitatively evaluates the importance of the smoothed step functions to the final result. Finally, for *Experiment 3*, the threshold for the normalized values of $p(\boldsymbol{\mathcal{D}}_i|\Delta\tilde{c}_{\mathrm{s}i}^+)$ and $p(\boldsymbol{\mathcal{D}}_i|\Delta\tilde{c}_{\mathrm{s}i}^-)$ is set to zero and we use a constant standard deviation of $\sigma_{\Delta\tilde{c}_{\mathrm{s}i}} = \sqrt{s_\mathrm{h}}$ instead of estimating it from $p(\boldsymbol{\mathcal{D}}_i|\Delta\tilde{c}_{\mathrm{s}i})$. Consequently, this disables the global approximation of first- and second-order partial derivatives of the log-likelihood and we are able to assess the performance using only local estimates for first-order partial derivatives.

The average success rates for the three experiments as well as for the unmodified tracker are given in Table 1. The final result highlights the importance of

Table 1. Average success rate for three experiments that feature different tracker configurations and for the *default* tracker. The evaluation is conducted on the *noise* sequence of the *RBOT* dataset, considering all 18 objects. *Experiment 1* shows the tracker without Tikhonov regularization, in *Experiment 2* a linear smoothed step function is used, for *Experiment 3* the approximation of first- and second-order partial derivatives was disabled, and results for the unmodified tracker are shown in the *default* column.

Experiment	1	2	3	default
Success Rate	46.8	59.7	65.1	71.5

each of the three tracker components that are evaluated in the ablation study. The low success rate of *Experiment 1* shows that the used Tikhonov regularization, which constrains the Newton optimization relative to the previous pose, is essential for the functioning of the tracker. Similarly, the results of *Experiment 2* provide a great example of the algorithm's bad performance if the smoothed step function is not designed to ensure Gaussian properties. Finally, while the results for *Experiment 3* are above the current state of the art, which achieves a success rate of 63.6, the experiment demonstrates that the tracker performs significantly better if global approximations for both the first- and second-order derivatives are used.