From images to localization

Contents

In class you have seen that given s, the outlier fraction ϵ and a desired success rate p, the required amount of iterations k can be calculated as follows:

$$k = \frac{\log(1-p)}{\log(1-(1-\epsilon)^s)} \tag{1}$$

1 Numerical Exercises

1. If the outlier ratio ϵ is not known before determing the number of iterations for RANSAC, it is possible to estimate the ratio ϵ adaptively while running RANSAC.

   (a) Derive the formula for the estimated outlier ratio $\hat{\epsilon}$. Hint: Try to find an upper bound on the outlier ratio, which is iteratively updated.
   Solution:
   To estimate the outlier ratio $\hat{\epsilon}$, we first compute a lower bound for the inlier ratio $\hat{w}_i$ at the current iteration step i, which is the fraction of the the number of inliers $n_{i,\text{max inliers}}$ corresponding to the largest set of inliers found up to the iteration step i over the total number of points n_{points}.

   $$\hat{w}_i = \frac{n_{i,\text{max inliers}}}{n_{\text{points}}}$$

   This estimated inlier ratio represents a lower bound since the current $n_{i,\text{max inliers}}$ is only updated if the number of inliers increases. Finally, the upper bound of the outlier ratio can be estimated from to the lower bound of the inlier ratio as follows.

   $$\hat{\epsilon}_i = 1 - \frac{n_{i,\text{max inliers}}}{n_{\text{points}}} \tag{2}$$

   (b) Is this way of estimating the outlier ratio guaranteed to consider the probability equation 1?
   Solution:
   Equation 1 holds if the ground truth outlier ratio ϵ_{gt} is correctly approximated by $\hat{\epsilon}$ in the corresponding k_{gt} number of iteration. If this is not the case, the effective number of iteration $\hat{k}$ can only be larger than the k_{gt} since we compute the number of iterations based on an upper bound of the outlier ratio. Thus, the estimated outlier ratio ensures that the correct model is selected with a probablity higher or equal to the propapility p_{gt} used in the case if the ground truth outlier ratio is known.

2. In the case of general model fitting, what are the possible failure cases for RANSAC?
   Solution:
   Possible failure cases are:

(a) There was never a subset consisting only of inliers selected during the predefined RANSAC iterations k, which is influced by the number of points s needed to fit the model, and the outlier ratio present in the data.

(b) There is a distribution modality in the data, which has more points fullfilling the model than the correct ground truth inliers. This can be also caused by a too small distance threshold.

(c) The distance function does not relate to the variables estimated by the model.

(d) The distance threshold to classify inliers is set too large such that outliers are detected as inliers.

3. Prove the convexity or non-convexity for the following two norms. You can use the the convexity condition for the second derivative for twice differentiable convex functions.

(a) Huber Norm
Solution:
To show that a function f is convex, it is sufficient to show that the second derivative is positive $f'' \geq 0$ over the function domain. To simplify the proof, we use the property that the derivative of even functions is odd and vice versa (the derivative of odd function is even). Based on this property, we know that the second derivative of the Huber norm is even since the Huber norm is even. Thus, it is enough to show that the second derivative is positive on the domain $[0, \infty)$.
The first and second derivative of the Huber norm are as follows:

$$\rho'(x) = \begin{cases} 2x & x \leq k \\ 2k & x > k \end{cases}$$

$$\rho''(x) = \begin{cases} 2 & x \leq k \\ 0 & x > k \end{cases}$$

To show that the second derivative exists at k, we need to prove that $\rho(x)$ and $\rho'(x)$ are continuous at k, which is the case:

$$\rho_+(x) = \lim_{x \to k^+} \rho(x) = k^2 = \lim_{x \to k^-} \rho(x) = \rho_-(x)$$

$$\rho'_+(x) = \lim_{x \to k^+} \rho'(x) = 2k = \lim_{x \to k^-} \rho'(x) = \rho'_-(x)$$

Thus, we have shown that the Huber norm is convex since the second derivative exists on the domain $[0, \infty)$ and is positive $\rho''(x) \geq 0$.

(b) Tukey Norm
Solution:
Following the same strategy as in (a), it is possible to show that the second derivative of the Tukey norm is not always positive $f'' \geq 0$ over the function domain. Instead, we can also consider the following condition, which needs to hold for any convex function f for all $0 < t < 1$ and all $x_1, x_2 \in X$.

$$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2)$$

We can show that the Tukey norm is not convex by proving that the above condition does not hold for $x_1 = 0$, $x_2 = 2\alpha$ and $t = 0.5$.

$$f(0 + \alpha) \nleq 0.5f(0) + 0.5f(2\alpha)$$

$$\alpha^2 \nleq 0 + 0.5\alpha^2$$

4. What is the probability of success p if n_{it} iterations are used for N datapoints containing n_{out} outlier datapoints and a minimum of s datapoints is required for estimating the model?
   Solution:

$$p = 1 - (1 - (1 - \frac{n_{out}}{N})^s)^{n_{it}}$$

5. Is the following statement true? If we have 1000 datapoints, a higher theoretical number of RANSAC iterations is required than with 100 datapoints to achieve the same probability of success. Assume that all datapoints are drawn from the same distribution.
   Solution: False