Day 10

Statistical Robustness

RANSAC

```
best model = {}
best consensus set = {}
best error = infinity
for i = 1..MAX ITERATIONS
  inlier set = randomly selected subset of data
  remainder set = data with inlier set removed
  consensus set = inlier set
  inlier model = fit inlier set to model
  for each element in remainder set
        if element fits inlier model then add it to consensus set
  end for
  if |consensus set| > MIN CONSENSUS SIZE
        consensus model = fit consensus set to model
        consensus error = residual error of consensus model
        if consensus error < best error
                best model = consensus model
                best consensus set = consensus set
                best error = consensus error
        end if
  end if
end for
```

LMS

```
best_model = {}
best_error = infinity
for each subset X of size p of the data
X_model = fit X to model (uses p data points)
X_error = median residual (over all N data points)
if X_error < best_error
        best_model = X_model
        best_error = X_error
end if</pre>
```

end for

Outlier Detection

- LMS produces robust but sub-optimal estimates
 - better estimates can be obtained by using LMS to obtain an estimate, removing the outlying data points, and then computing a least-squares estimates from the cleaned data set
- outliers can be identified using a priori knowledge (if you have it)
 - e.g., using Horn's method with a typical optical tracking system it would be reasonable to expect inliers to have an FRE < 1mm
- can also obtain a robust estimate of scale and label a data point as an outlier if its squared residual error is greater than some constant times the squared scale estimate
 - P.J. Rousseeuw and A.M. Leroy. Robust Regression and Outlier Detection. John Wiley & Sons, New York, 1987.