RANSAC
RANdom SAmple Consensus

Tomáš Svoboda, svoboda@cmp.felk.cvut.cz
courtesy of Ondřej Chum, Jiří Matas
Czech Technical University in Prague, Faculty of Electrical Engineering, Department of Cybernetics, Center for Machine Perception
http://cmp.felk.cvut.cz
Last update: November 20, 2009;

◆ importance for robust model estimation
◆ principle
◆ application
Importance for Computer Vision

- one of the most cited papers in computer vision and related fields (around 3900 citations according to Google scholar in 11/2009)
- widely accepted as a method that works even for very difficult problems
- recent advancement presented at the “25-years of RANSAC” workshop\(^1\). Look at the R. Bowless’ presentation.

\(^1\)http://cmp.felk.cvut.cz/ransac-cvpr2006
LS does not work for gross errors . . .

\(^2\)sketch borrowed from [3]
RANSAC motivations

- gross errors (outliers) spoil LS estimation
- detection (localization) algorithms in computer vision and recognition do have gross error
- in difficult problems the portion of good data may be even less than $\frac{1}{2}$
- standard robust estimation techniques [5] hardly applicable to data with less than $\frac{1}{2}$ “good” samples (points, lines, . . .)
RANSAC inputs and output

**In:**
- \( U = \{x_i\} \) set of data points, \(|U| = N\)
- \( f(S) : S \rightarrow \theta \) function \( f \) computes model parameters \( \theta \) given a sample \( S \) from \( U \)
- \( \rho(\theta, x) \) the cost function for a single data point \( x \)

**Out:**
- \( \theta^* \) \( \theta^* \), parameters of the model maximizing (or minimizing) the cost function
RANSAC inputs and output

**In:**

- $U = \{x_i\}$ set of data points, $|U| = N$
- $f(S) : S \rightarrow \theta$ function $f$ computes model parameters $\theta$
given a sample $S$ from $U$
- $\rho(\theta, x)$ the cost function for a single data point $x$

**Out:**

- $\theta^*$ $\theta^*$, parameters of the model maximizing (or minimizing) the cost function

RANSAC principle

1. **select randomly** few samples needed for model estimation
2. **verify** the model
3. **keep** the best so far model estimated
4. if **enough** trials then quit otherways repeat
RANSAC algorithm

\[ k := 0 \]

Repeat until \( P\{\text{better solution exists}\} < \eta \) (a function of \( C^* \) and no. of steps \( k \) )

\[ k := k + 1 \]

I. Hypothesis

(1) select randomly set \( S_k \subset U \), \( |S_k| = s \)

(2) compute parameters \( \theta_k = f(S_k) \)

II. Verification

(3) compute cost \( C_k = \sum_{x \in U} \rho(\theta_k, x) \)

(4) if \( C^* < C_k \) then \( C^* := C_k, \ \theta^* := \theta_k \)

end
Explanation example: line detection
Explanation example: line detection

- Randomly select two points
Explanation example: line detection

- Randomly select two points
- The hypothesised model is the line passing through the two points
Explanation example: line detection

- Randomly select two points
- The hypothesised model is the line passing through the two points
- The error function is a distance from the line
Explanation example: line detection

- Randomly select two points
- The hypothesised model is the line passing through the two points
- The error function is a distance from the line
- Points consistent with the model
Probability of selecting uncontaminated sample in $K$ trials

- $N$ - number of data points
- $w$ - fraction of inliers
- $s$ - size of the sample
Probability of selecting uncontaminated sample in \( K \) trials

- \( N \) - number of data points
- \( w \) - fraction of inliers
- \( s \) - size of the sample

Prob. of selecting a sample with all inliers:\(^3\)
Probability of selecting uncontaminated sample in $K$ trials

- $N$ - number of data points
- $w$ - fraction of inliers
- $s$ - size of the sample

Prob. of selecting a sample with all inliers$^3$: $\approx w^s$
Probability of selecting uncontaminated sample in $K$ trials

- $N$ - number of data points
- $w$ - fraction of inliers
- $s$ - size of the sample

Prob. of selecting a sample with all inliers$^3$: $\approx w^s$

Prob. of not selecting a sample with all inliers:
Probability of selecting uncontaminated sample in $K$ trials

- $N$ - number of data points
- $w$ - fraction of inliers
- $s$ - size of the sample

Prob. of selecting a sample with all inliers $^3$: $\approx w^s$
Prob. of not selecting a sample with all inliers: $1 - w^s$
Probability of selecting uncontaminated sample in $K$ trials

- $N$ - number of data points
- $w$ - fraction of inliers
- $s$ - size of the sample

Prob. of selecting a sample with all inliers$^3$: $\approx w^s$

Prob. of not selecting a sample with all inliers: $1 - w^s$

Prob. of not selecting a good sample $K$ times:
Probability of selecting uncontaminated sample in $K$ trials

- $N$ - number of data points
- $w$ - fraction of inliers
- $s$ - size of the sample

Prob. of selecting a sample with all inliers: $\approx w^s$

Prob. of not selecting a sample with all inliers: $1 - w^s$

Prob. of not selecting a good sample $K$ times: $(1 - w^s)^K$
Probability of selecting uncontaminated sample in $K$ trials

- $N$ - number of data points
- $w$ - fraction of inliers
- $s$ - size of the sample

Prob. of selecting a sample with all inliers $^3$: $\approx w^s$

Prob. of not selecting a sample with all inliers: $1 - w^s$

Prob. of not selecting a good sample $K$ times: $(1 - w^s)^K$

Prob. of selecting uncontaminated sample in $K$ trials at least once:
Probability of selecting uncontaminated sample in $K$ trials

- $N$ - number of data points
- $w$ - fraction of inliers
- $s$ - size of the sample

Prob. of selecting a sample with all inliers\(^3\): $\approx w^s$
Prob. of not selecting a sample with all inliers: $1 - w^s$
Prob. of not selecting a good sample $K$ times: $(1 - w^s)^K$

Prob. of selecting uncontaminated sample in $K$ trials at least once:
$$P = 1 - (1 - w^s)^K$$

\(^3\)Approximation valid for $s \ll N$, see the lecture notes
How many samples are needed, $K = \ ?$

How many trials is needed to select an uncontaminated sample with a given probability $P$? We derived $P = 1 - (1 - w^s)^K$. Log the both sides to get

$$K = \frac{\log(1 - P)}{\log(1 - w^s)}$$

![Graph showing the number of samples needed for different sample sizes and probabilities]
Real problem—\( w \) unknown

Often, the proportion of inliers in data cannot be estimated in advance.

**Adaptive estimation:** start with worst case and update the estimate as the computation progress

- set \( K = \infty \), \#samples = 0, \( P \) very conservative, say \( P = 0.99 \)
- while \( K > \#samples \) repeat
  - choose a random sample, compute the model and count inliers
  - \( w = \frac{\#\text{inliers}}{\#\text{data points}} \)
  - \( K = \frac{\log(1-P)}{\log(1-w^s)} \)
  - increment \#samples
- terminate
Fitting line via RANSAC

![Image](http://visionbook.felk.cvut.cz)

video:fitting_line
Epipolar geometry estimation by RANSAC

- $U$: a set of correspondences, i.e. pairs of 2D points
- $s = 7$
- $f$: seven-point algorithm - gives 1 to 3 independent solutions
- $\rho$: thresholded Sampson’s error
References

Besides the main reference [2] the Huber’s book [5] about robust estimation is also widely recognized. The RANSAC algorithm recieved several essential improvements in recent years [1, 6, 7].

For the seven-point algorithm and Sampson’s error, see [4]


End
Prob. of selecting at least one uncontaminated sample $p=0.99$

<table>
<thead>
<tr>
<th>sample size</th>
<th>number of samples needed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$w=0.5$

$w=0.7$