RANSAC
RANdom SAmple Consensus

Tomáš Svoboda, svoboda@cmp.felk.cvut.cz
courtesy of Ondřej Chum, Jiří Matas
Czech Technical University in Prague, Center for Machine Perception
http://cmp.felk.cvut.cz
Last update: October 27, 2008;

Talk Outline

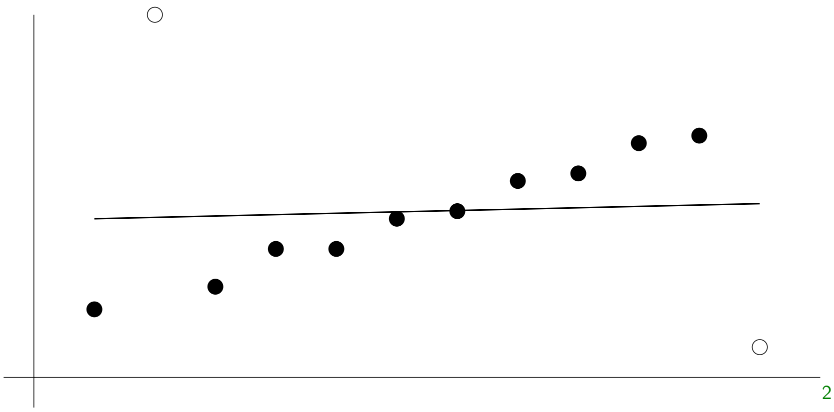
- importance for computer vision
- principle
- line fitting
- epipolar geometry

Importance for Computer Vision

- published in 1981 as a model fitting method [2]
- on of the most cited papers in computer vision and related fields
- widely accepted as a method that works even for difficult computer vision 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

LSQ does not work for gross errors . . .


[2]

[2] sketch borrowed from [3]

RANSAC motivations for computer vision

- gross errors (outliers) spoil LSQ 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 $1/2$
- standard robust estimation techniques hardly applicable to data with less than $1/2$ good

RANSAC inputs and output

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

RANSAC algorithm

$k := 0$

Repeat until P{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 $\theta_k = f(S_k)$

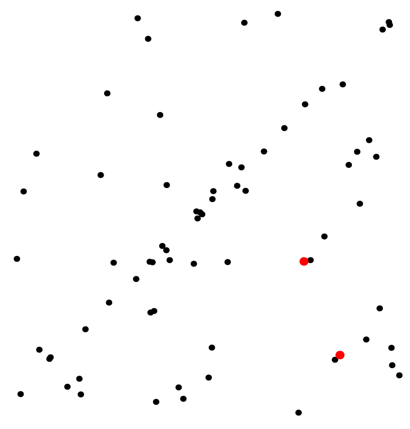
II. Verification

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

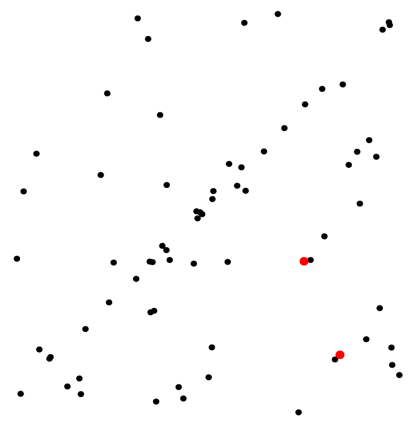
(4) if $C^* < C_k$ then $C^* := C_k$, $\theta^* := \theta_k$

end

Explanation example: line detection

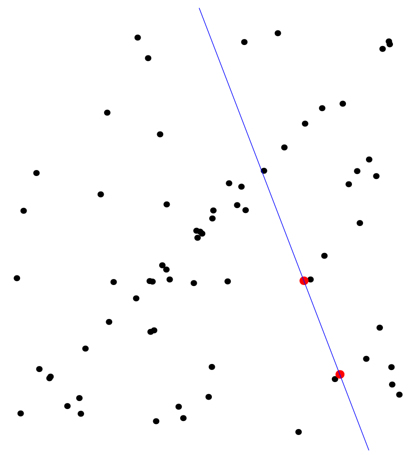


Explanation example: line detection



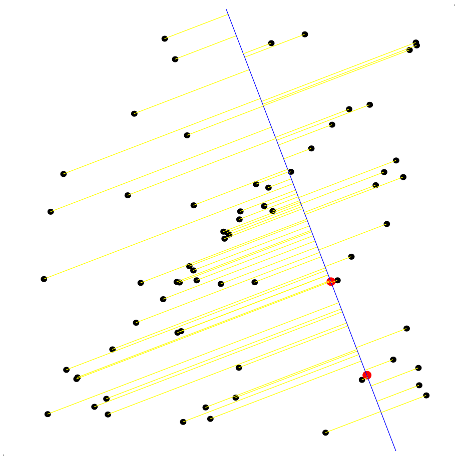
- Randomly select two points

Explanation example: line detection



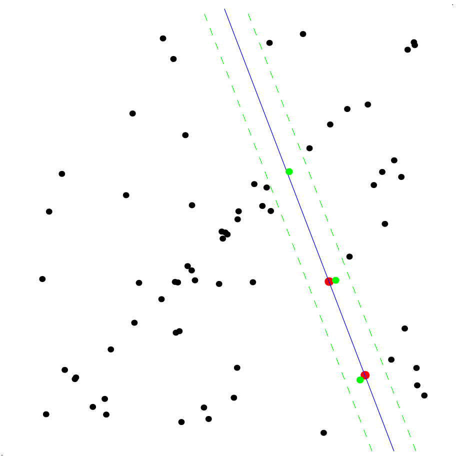
- Randomly select two points
- The hypothesised model is the line passing through the two points

Explanation example: line detection



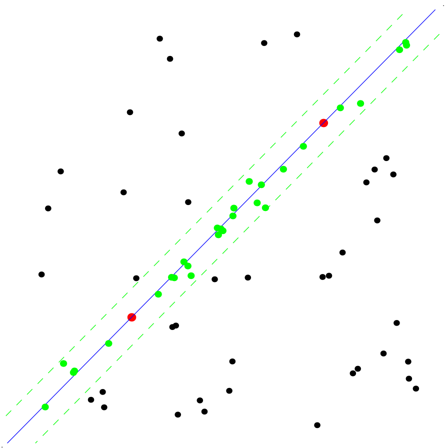
- Randomly select two points
- The hypothesised model is the line passing through the two points
- The error function is a distance from the line

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



Uncontaminated sample

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

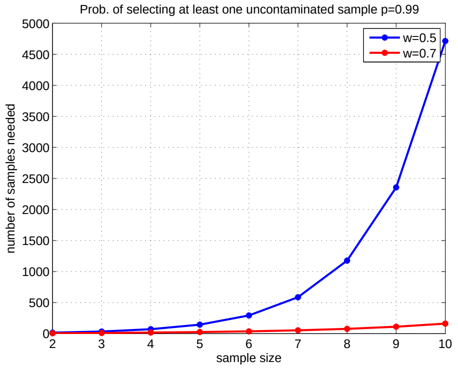
Prob. of selecting a sample with all inliers[3]: $\approx w^s$
Prob. of not selecting a sample with all inliers: $1 - w^s$
Prob. of not selecting a good sample K times: $(1 - w^s)^K$
The sought probability of selecting uncontaminated sample in K trials at least once: $P = 1 - (1 - w^s)^K$

[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

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



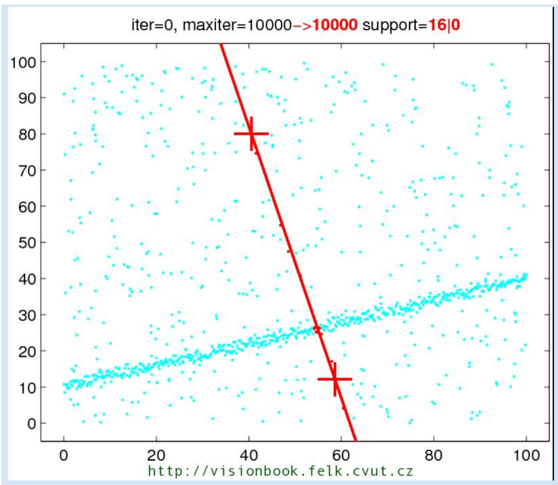
Real problem—w unknown

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

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

- set $K = \infty$, $\#\text{samples} = 0$, P very conservative, say $P = 0.99$
- while $K > \#\text{samples}$ 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 $\#\text{samples}$
- terminate

Fitting line via RANSAC



video:fitting line

Epipolar geometry estimation by RANSAC

- U : a set of correspondences, i.e. pairs of 2D points — data points
- $s = 7$ — sample size
- f : seven-point algorithm - gives 1 to 3 independent solutions — model parameters
- ρ : thresholded Sampson's error — cost function

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]

[1] Ondřej Chum and Jiří Matas. Matching with PROSAC - progressive sample consensus. In Cordelia Schmid, Stefano Soatto, and Carlo Tomasi, editors, Proc. of Conference on Computer Vision and Pattern Recognition (CVPR), volume 1, pages 220–226, Los Alamitos, USA, June 2005. IEEE Computer Society.

[2] M.A. Fischler and R.C. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM, 24(6):381–395, June 1981.

[3] R. Hartley and A. Zisserman. Multiple View Geometry in Computer Vision. Cambridge University Press, Cambridge, UK, 2000. On-line resources at: http://www.robots.ox.ac.uk/~vgg/hzbook/hzbook1.html.

[4] Richard Hartley and Andrew Zisserman. Multiple view geometry in computer vision. Cambridge University, Cambridge, 2nd edition, 2003.

[5] Peter J. Huber. Robust Statistics. Willey series in probability and mathematical statistics. John Willey and Sons, 1981.

[6] Jiří Matas and Ondřej Chum. Randomized RANSAC with $T_{d,d}$ test. Image and Vision Computing, 22(10):837–842, September 2004.

[7] Jiří Matas and Ondřej Chum. Randomized ransac with sequential probability ratio test. In Songde Ma and Heung-Yeung Shum, editors, Proc. IEEE International Conference on Computer Vision (ICCV), volume II, pages 1727–1732, New York, USA, October 2005. IEEE Computer Society Press.

End