Deep Learning for Computer Vision

Feature Matching

Vineeth N Balasubramanian

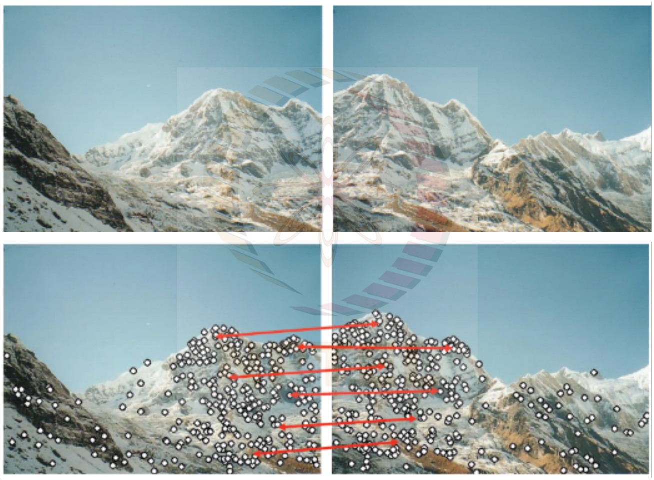
Department of Computer Science and Engineering
Indian Institute of Technology, Hyderabad

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Acknowledgements

- Most of this lecture's slides are based on lectures of **Deep Learning for Vision** course taught by Prof Yannis Avrithis at Inria Rennes-Bretagne Atlantique

Review



How to match?

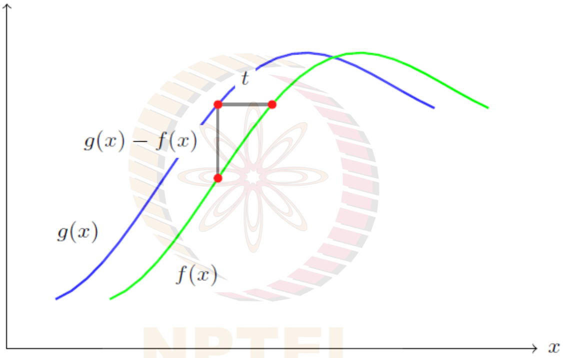
Dense Registration through Optical Flow[1]

- For each location in an image, find a displacement with respect to another reference image
- Appropriate for small displacements, e.g. stereopsis or optical flow

[1] Lucas and Kanade IJCAI 1981. An Iterative Image Registration Technique With an Application to Stereo Vision.

Dense Registration through Optical Flow[2]

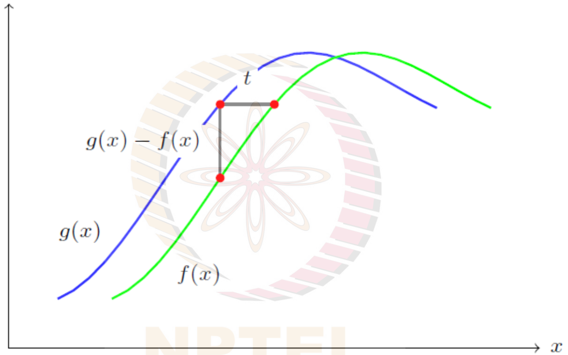
- One dimension:



[2] Lucas and Kanade IJCAI 1981. An Iterative Image Registration Technique With an Application to Stereo Vision.

Dense Registration through Optical Flow[2]

- One dimension:



Assuming $g(x) = f(x+t)$ and t is small,

$$\frac{df}{dx}(x) \approx \frac{f(x+t) - f(x)}{t} = \frac{g(x) - f(x)}{t}$$

[2]Lucas and Kanade IJCAI 1981. An Iterative Image Registration Technique With an Application to Stereo Vision.

Dense Registration through Optical Flow[2]

- Error given by:

$$E(t) = \sum_x w(x)\Big(f(x+t) - g(x)\Big)^2 \approx \sum_x w(x)\Big(f(x) + t^T \Delta f(x) - g(x)\Big)^2$$

[2] Lucas and Kanade IJCAI 1981. An Iterative Image Registration Technique With an Application to Stereo Vision.

Dense Registration through Optical Flow[2]

- Error given by:

$$E(t) = \sum_x w(x)\Big(f(x+t) - g(x)\Big)^2 \approx \sum_x w(x)\Big(f(x) + t^T \Delta f(x) - g(x)\Big)^2$$

- Error minimized when gradient vanishes

$$\frac{\partial E}{\partial t} = \sum_x w(x) 2\Delta f(x)\Big(f(x) + t^T \Delta f(x) - g(x)\Big) = 0$$

[2] Lucas and Kanade IJCAI 1981. An Iterative Image Registration Technique With an Application to Stereo Vision.

Dense Registration through Optical Flow[2]

- Error given by:

$$E(t) = \sum_x w(x)\Big(f(x+t) - g(x)\Big)^2 \approx \sum_x w(x)\Big(f(x) + t^T \Delta f(x) - g(x)\Big)^2$$

- Error minimized when gradient vanishes

$$\frac{\partial E}{\partial t} = \sum_x w(x) 2\Delta f(x)\Big(f(x) + t^T \Delta f(x) - g(x)\Big) = 0$$

- Least-squares solution *(ignoring summation and arguments for simplicity)*:

$$w\Delta f(\Delta f)^T t = w\Delta f(g - f)$$

[2] Lucas and Kanade IJCAI 1981. An Iterative Image Registration Technique With an Application to Stereo Vision.

Dense Registration through Optical Flow[2]

- Error given by:

$$E(t) = \sum_x w(x)\Big(f(x+t) - g(x)\Big)^2 \approx \sum_x w(x)\Big(f(x) + t^T \Delta f(x) - g(x)\Big)^2$$

- Error minimized when gradient vanishes

$$\frac{\partial E}{\partial t} = \sum_x w(x) 2\Delta f(x)\Big(f(x) + t^T \Delta f(x) - g(x)\Big) = 0$$

- Least-squares solution *(ignoring summation and arguments for simplicity)*:

$$w\Delta f(\Delta f)^T t = w\Delta f(g - f)$$

- 2-D equivalent: Assume an image patch defined by window w; what is the error between patch shifted by t in reference image f and patch at origin in shifted image g?

[2] Lucas and Kanade IJCAI 1981. An Iterative Image Registration Technique With an Application to Stereo Vision.

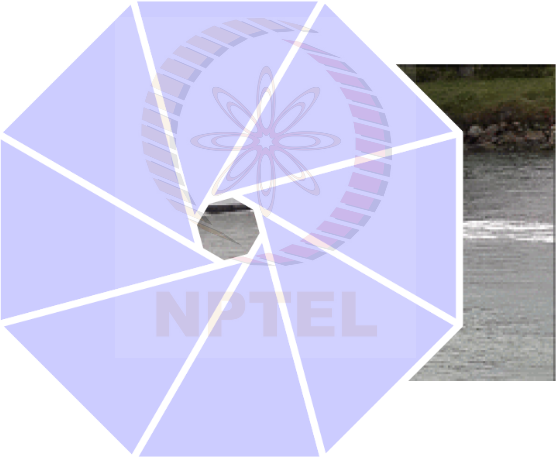
Dense Registration through Optical Flow[2]

- The Aperture Problem:

[2] Lucas and Kanade IJCAI 1981. An Iterative Image Registration Technique With an Application to Stereo Vision.

Dense Registration through Optical Flow[2]

- The Aperture Problem:



[2]Lucas and Kanade IJCAI 1981. An Iterative Image Registration Technique With an Application to Stereo Vision.

Wide Baseline Spatial Matching

- In dense registration, we started from a local “template matching” process and found an efficient solution based on a Taylor approximation
- Both make sense for small displacements
- In wide-baseline matching, every part of one image may appear anywhere in the other
- We start by pairwise matching of local descriptors without any order, and then attempt to enforce some geometric consistency according to a rigid motion model

Wide Baseline Spatial Matching

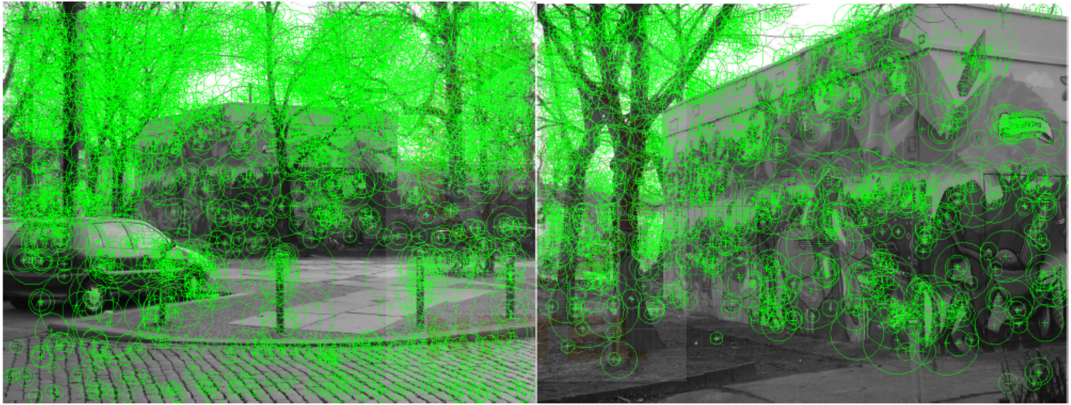
- In dense registration, we started from a local "template matching" process and found an efficient solution based on a Taylor approximation
- Both make sense for small displacements
- In wide-baseline matching, every part of one image may appear anywhere in the other
- We start by pairwise matching of local descriptors without any order, and then attempt to enforce some geometric consistency according to a rigid motion model

Wide Baseline Spatial Matching

A region in one image may appear anywhere in the other

Credit: Yannis Avrithis, Inria Rennes-Bretagne Atlantique

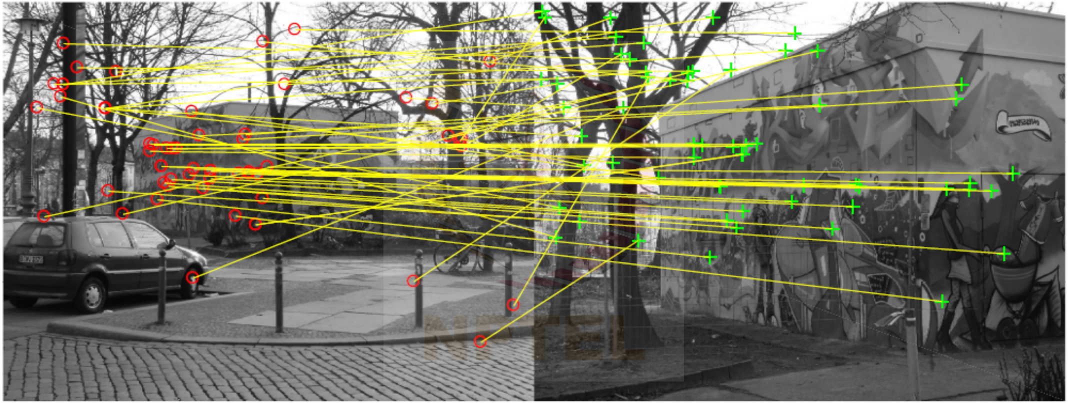
Wide Baseline Spatial Matching



Features detected independently in each image

Credit: Yannis Avrithis, Inria Rennes-Bretagne Atlantique

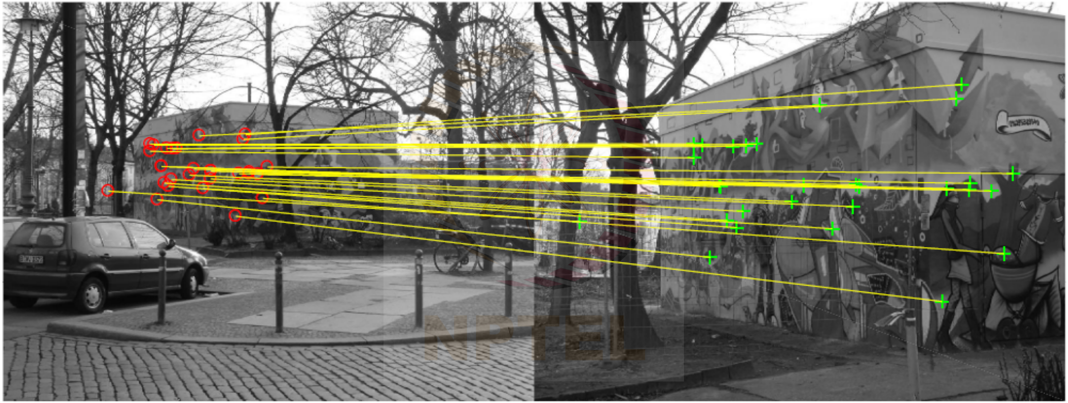
Wide Baseline Spatial Matching



Tentative correspondences by pairwise descriptor matching

Credit: Yannis Avrithis, Inria Rennes-Bretagne Atlantique

Wide Baseline Spatial Matching



Subset of correspondences that are ‘inlier’ to a rigid transformation

Credit: Yannis Avrithis, Inria Rennes-Bretagne Atlantique

Wide Baseline Spatial Matching

Descriptor Extraction:

For each detected feature in each image:

- Construct a local histogram of gradient orientations (HoG)
- Find one or more dominant orientations corresponding to peaks in histogram
- Resample local patch at given location, scale, and orientation
- Extract one descriptor for each dominant orientation

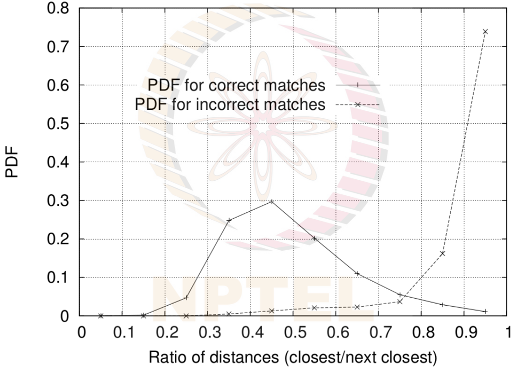
Wide Baseline Spatial Matching

Descriptor Matching:

- For each descriptor in one image, find its two nearest neighbors in the other
- If ratio of distance of first to distance of second is small, make a correspondence
- This yields a list of tentative correspondences

Wide Baseline Spatial Matching

Ratio Test:



Ratio of first to second nearest neighbour distance can determine the probability of a true correspondence

Credit: Yannis Avrithis, Inria Rennes-Bretagne Atlantique

Wide Baseline Spatial Matching

Why is it difficult?

- Should allow for a geometric transformation
- Fitting the model to data (correspondences) is sensitive to outliers: should find a subset of inliers first
- Finding inliers to a transformation requires finding the transformation in the first place
- Correspondences can have gross error
- Inliers are typically less than 50%

Geometric Transformations

- Two images I, I' are equal at points $\mathbf{x}, \mathbf{x}'$

$$I(\mathbf{x}) = I'(\mathbf{x}')$$

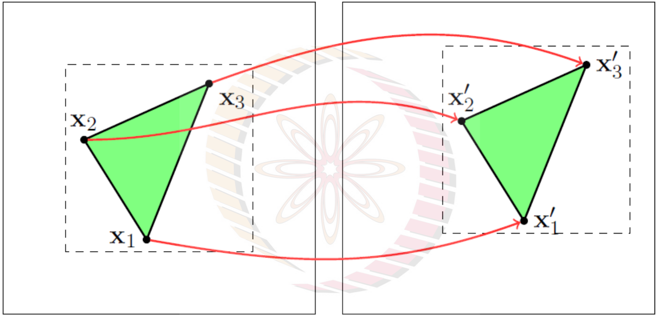
- $\mathbf{x}$ is mapped to $\mathbf{x}'$

$$\mathbf{x}' = T(\mathbf{x})$$

- T is a bijection of $\mathbb{R}^2$ to itself:

$$T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$$

Geometric Transformations

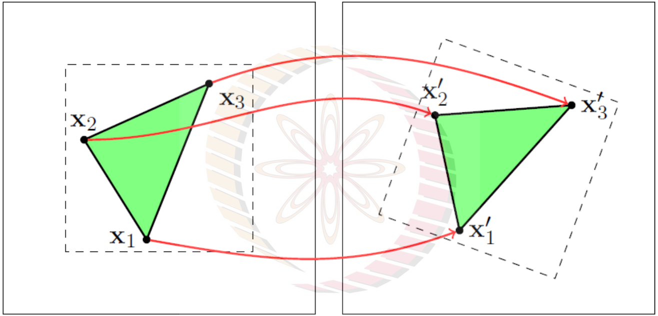


- Translation: 2 degrees of freedom

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

Credit: Yannis Avrithis, Inria Rennes-Bretagne Atlantique

Geometric Transformations

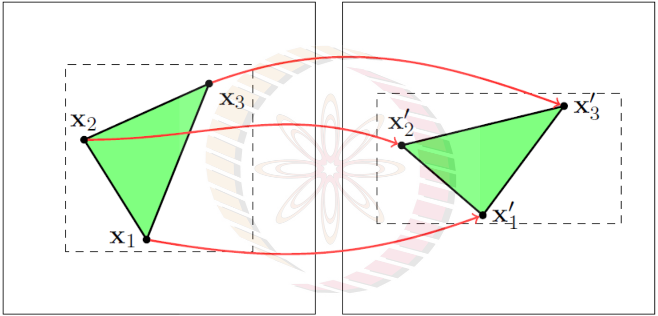


- Rotation: 1 degree of freedom

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

Credit: Yannis Avrithis, Inria Rennes-Bretagne Atlantique

Geometric Transformations

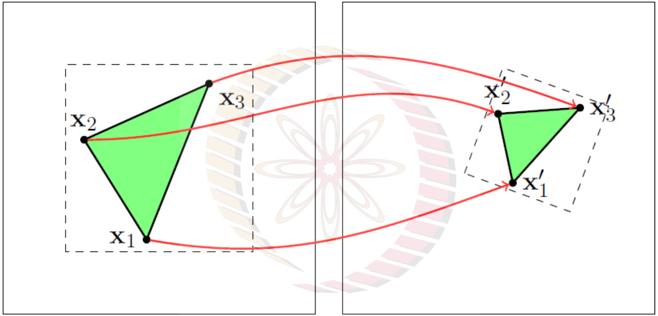


- Similarity: 4 degrees of freedom

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} r\cos\theta & -r\sin\theta & t_x \\ r\sin\theta & r\cos\theta & t_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

Credit: Yannis Avrithis, Inria Rennes-Bretagne Atlantique

Geometric Transformations

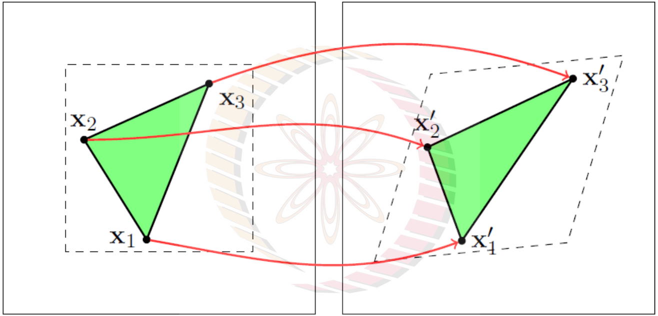


- Similarity: 4 degrees of freedom

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} r\cos\theta & -r\sin\theta & t_x \\ r\sin\theta & r\cos\theta & t_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

Credit: Yannis Avrithis, Inria Rennes-Bretagne Atlantique

Geometric Transformations

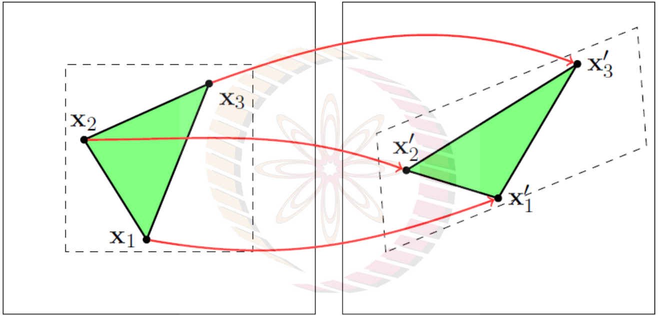


- Shear: 2 degrees of freedom

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & b_x & 0 \\ b_y & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

Credit: Yannis Avrithis, Inria Rennes-Bretagne Atlantique

Geometric Transformations



- Affine: 6 degrees of freedom

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

Credit: Yannis Avrithis, Inria Rennes-Bretagne Atlantique

Correspondence and Least Squares

- In all cases, the problem is transformed to a linear system (why?)

$$A\mathbf{x} = \mathbf{b}$$

  where $\mathbf{x}$, $\mathbf{b}$ contain coordinates of known point correspondences from images I, I' respectively, and A contains our model parameters

Correspondence and Least Squares

- In all cases, the problem is transformed to a linear system (why?)

$$A\mathbf{x} = \mathbf{b}$$

  where $\mathbf{x}$, $\mathbf{b}$ contain coordinates of known point correspondences from images I, I' respectively, and A contains our model parameters

- We need $n = \lceil d/2 \rceil$ correspondences, where d are the degrees of freedom of our model

Correspondence and Least Squares

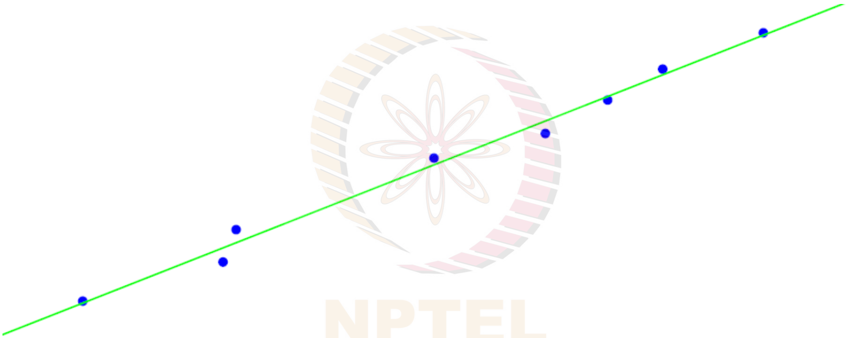
- In all cases, the problem is transformed to a linear system (why?)

$$A\mathbf{x} = \mathbf{b}$$

  where $\mathbf{x}$, $\mathbf{b}$ contain coordinates of known point correspondences from images I, I' respectively, and A contains our model parameters

- We need $n = \lceil d/2 \rceil$ correspondences, where d are the degrees of freedom of our model
- Let's take the simplest model as an example: fit a line to two points

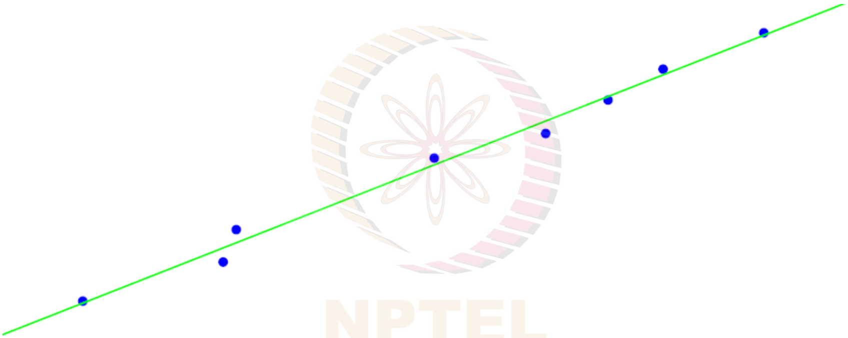
Correspondence and Least Squares



- clean data, no outliers : least squares fit ok

Credit: Yannis Avrithis, Inria Rennes-Bretagne Atlantique

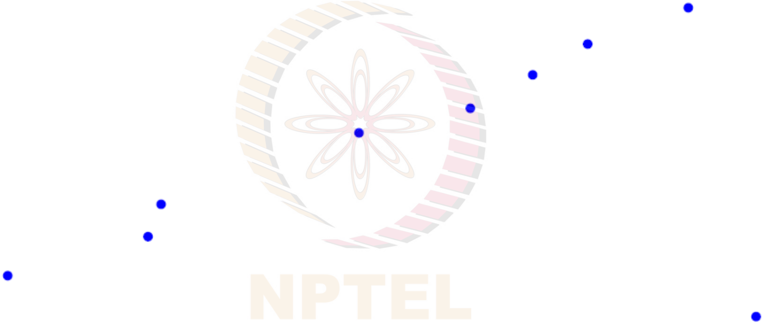
Correspondence and Least Squares



- clean data, no outliers : least squares fit ok

Credit: Yannis Avrithis, Inria Rennes-Bretagne Atlantique

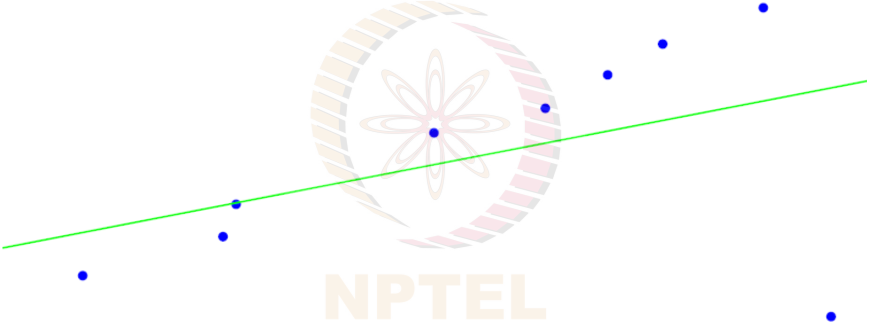
Correspondence and Least Squares



- one gross outlier - least squares fit fails - what do we do?

Credit: Yannis Avrithis, Inria Rennes-Bretagne Atlantique

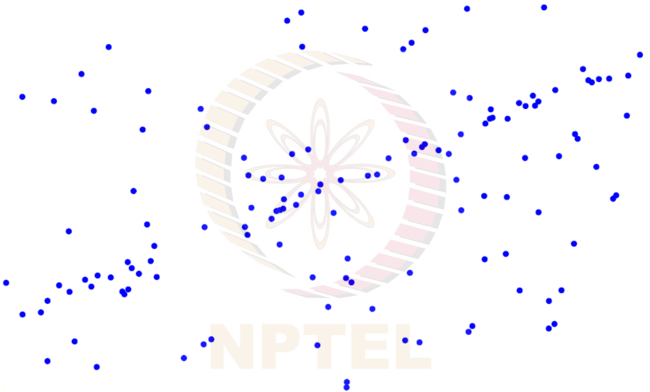
Correspondence and Least Squares



- one gross outlier - least squares fit fails - what do we do?

Credit: Yannis Avrithis, Inria Rennes-Bretagne Atlantique

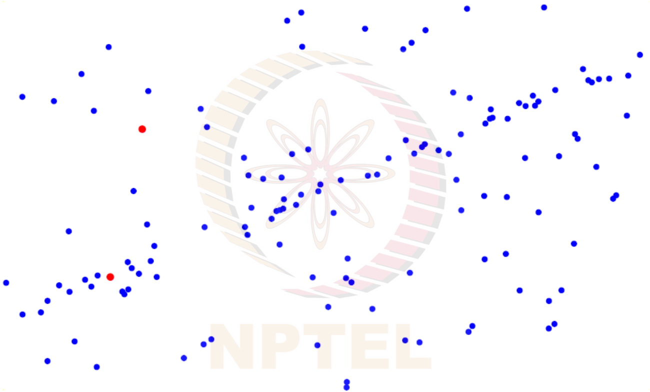
RANSAC (**RAN**dom **SA**mple **C**onsensus)[3]



- data with outliers - pick two points at random - draw line through them - set margin on either side - count inlier points

[3]Fischler and Bolles. CACM 1981. Random Sample Consensus: A Paradigm for Model Fitting With Applications to Image Analysis and Automated Cartography.

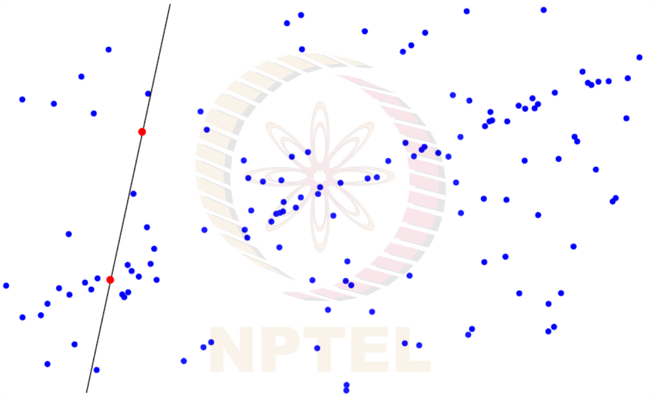
RANSAC (**RAN**dom **SA**mple **C**onsensus)[3]



- data with outliers - pick two points at random - draw line through them - set margin on either side - count inlier points

[3] Fischler and Bolles. CACM 1981. Random Sample Consensus: A Paradigm for Model Fitting With Applications to Image Analysis and Automated Cartography.

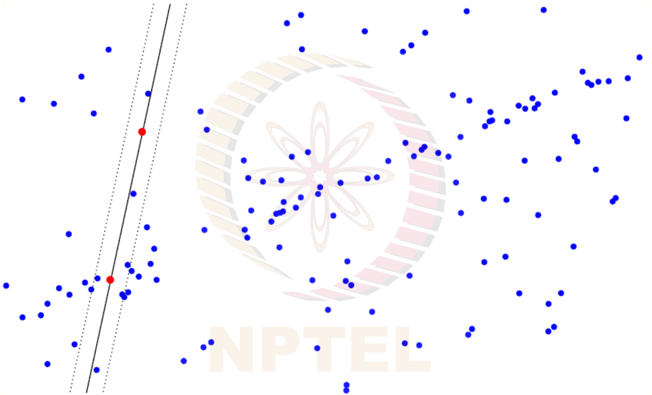
RANSAC (RANdom SAmple Consensus)[3]



- data with outliers - pick two points at random - draw line through them - set margin on either side - count inlier points

[3] Fischler and Bolles. CACM 1981. Random Sample Consensus: A Paradigm for Model Fitting With Applications to Image Analysis and Automated Cartography.

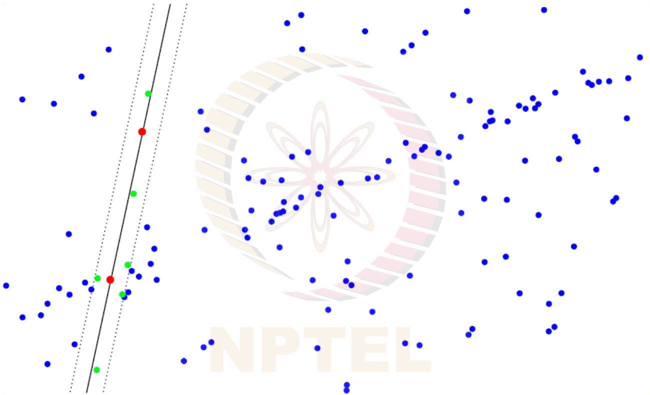
RANSAC (**RAN**dom **SA**mple **C**onsensus)[3]



- data with outliers - pick two points at random - draw line through them - set margin on either side - count inlier points

[3] Fischler and Bolles. CACM 1981. Random Sample Consensus: A Paradigm for Model Fitting With Applications to Image Analysis and Automated Cartography.

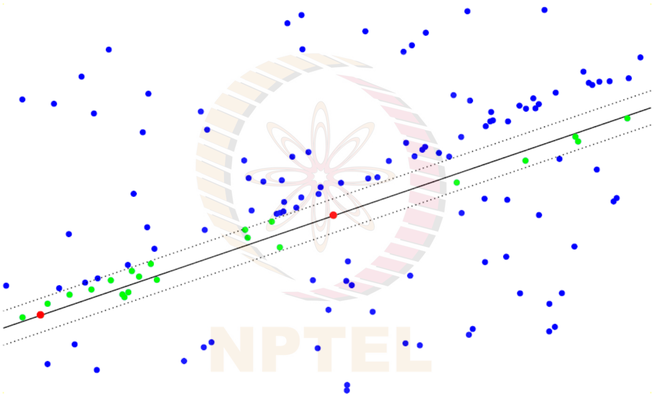
RANSAC (RANdom SAmple Consensus)[3]



- data with outliers - pick two points at random - draw line through them - set margin on either side - count inlier points

[3]Fischler and Bolles. CACM 1981. Random Sample Consensus: A Paradigm for Model Fitting With Applications to Image Analysis and Automated Cartography.

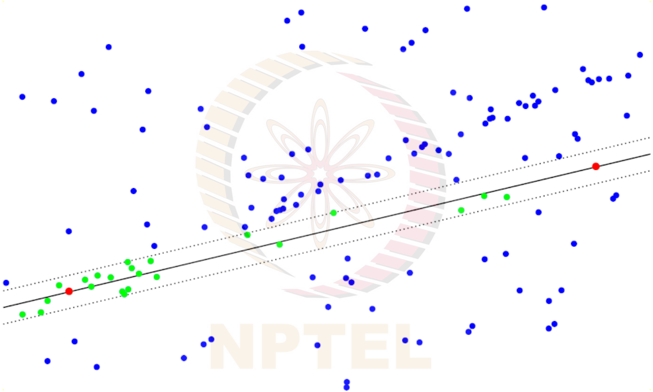
RANSAC (RANdom SAmple Consensus)[3]



- Repeat: pick two points at random, draw line through them, count inlier points at fixed distance to line, keep best hypothesis so far

[3]Fischler and Bolles. CACM 1981. Random Sample Consensus: A Paradigm for Model Fitting With Applications to Image Analysis and Automated Cartography.

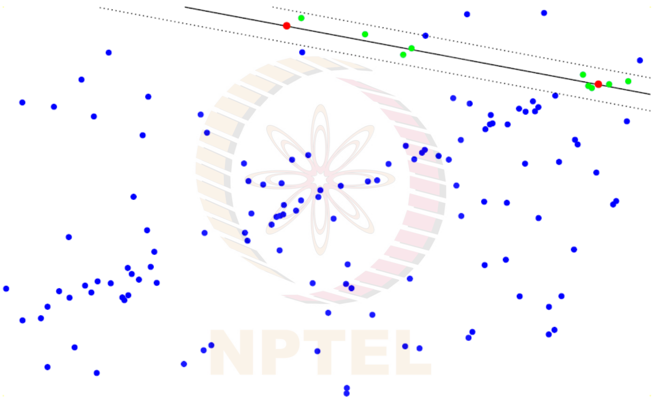
RANSAC (**RAN**dom **SA**mple **C**onsensus)[3]



- Repeat: pick two points at random, draw line through them, count inlier points at fixed distance to line, keep best hypothesis so far

[3]Fischler and Bolles. CACM 1981. Random Sample Consensus: A Paradigm for Model Fitting With Applications to Image Analysis and Automated Cartography.

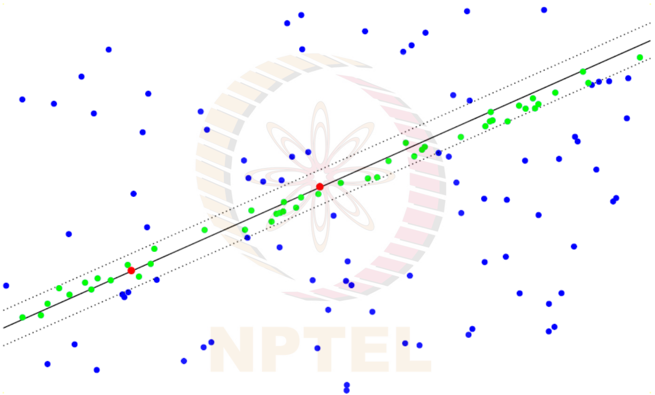
RANSAC (RANdom SAmple Consensus)[3]



- Repeat: pick two points at random, draw line through them, count inlier points at fixed distance to line, keep best hypothesis so far

[3] Fischler and Bolles. CACM 1981. Random Sample Consensus: A Paradigm for Model Fitting With Applications to Image Analysis and Automated Cartography.

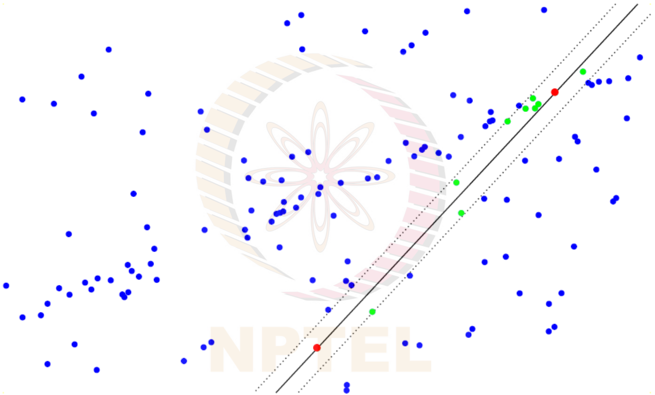
RANSAC (**RAN**dom **SA**mple **C**onsensus)[3]



- Repeat: pick two points at random, draw line through them, count inlier points at fixed distance to line, keep best hypothesis so far

[3] Fischler and Bolles. CACM 1981. Random Sample Consensus: A Paradigm for Model Fitting With Applications to Image Analysis and Automated Cartography.

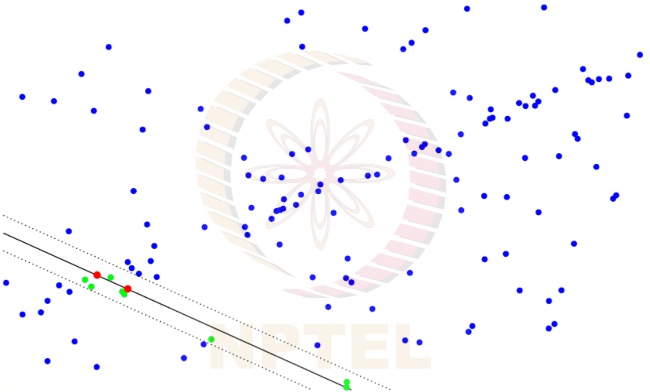
RANSAC (**RAN**dom **SA**mple **C**onsensus)[3]



- Repeat: pick two points at random, draw line through them, count inlier points at fixed distance to line, keep best hypothesis so far

[3]Fischler and Bolles. CACM 1981. Random Sample Consensus: A Paradigm for Model Fitting With Applications to Image Analysis and Automated Cartography.

RANSAC (**RAN**dom **SA**mple **C**onsensus)[3]



- Repeat: pick two points at random, draw line through them, count inlier points at fixed distance to line, keep best hypothesis so far

[3]Fischler and Bolles. CACM 1981. Random Sample Consensus: A Paradigm for Model Fitting With Applications to Image Analysis and Automated Cartography.

RANSAC

- X: data (tentative correspondences)
- n: minimum number of samples to fit a model
- $s(x;\theta)$: score of sample $\mathbf{x}$ given model parameters θ
- repeat:
  - hypothesis
    - draw n samples $H \subset X$ at random
    - fit model to H, compute parameters θ
  - verification
    - are data consistent with hypothesis? compute score $S = \sum_{\mathbf{x} \in X} s(\mathbf{x};\theta)$
    - if $S^* > S$, store solution $\theta^* := \theta$, $S^* := S$

RANSAC: Limitations

- Inlier ratio w (number of inliers in data / number of points in data) unknown
- Too expensive when minimum number of samples is large (e.g. $n > 6$) and inlier ratio is small (e.g. $w < 10\%$): 10^6 iterations for 1% probability of failure. (How?)

RANSAC: Limitations

- Inlier ratio w (number of inliers in data / number of points in data) unknown
- Too expensive when minimum number of samples is large (e.g. $n > 6$) and inlier ratio is small (e.g. $w < 10\%$): 10^6 iterations for 1% probability of failure. (How?)
    - $w^n \rightarrow$ probability that all n points are inliers

RANSAC: Limitations

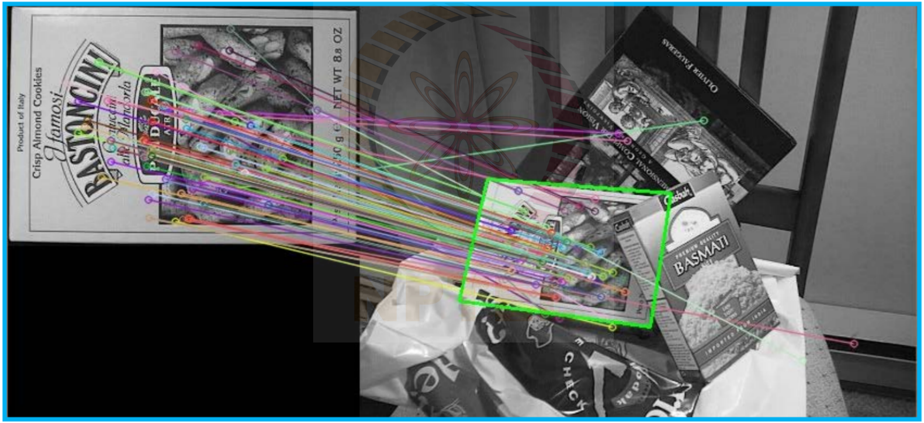
- Inlier ratio w (number of inliers in data / number of points in data) unknown
- Too expensive when minimum number of samples is large (e.g. $n > 6$) and inlier ratio is small (e.g. $w < 10\%$): 10^6 iterations for 1% probability of failure. (How?)
  - $w^n \rightarrow$ probability that all n points are inliers
  - $1 - w^n \rightarrow$ probability that at least one of n points is an outlier $\implies$ a bad model will be estimated from this point set

RANSAC: Limitations

- Inlier ratio w (number of inliers in data / number of points in data) unknown
- Too expensive when minimum number of samples is large (e.g. $n > 6$) and inlier ratio is small (e.g. $w < 10\%$): 10^6 iterations for 1% probability of failure. (How?)
  - $w^n \rightarrow$ probability that all n points are inliers
  - $1 - w^n \rightarrow$ probability that at least one of n points is an outlier $\implies$ a bad model will be estimated from this point set
  - $(1 - w^n)^k \rightarrow$ probability that algorithm never selects a set of n points which all are inliers, where $k \rightarrow$ number of iterations

RANSAC Applications

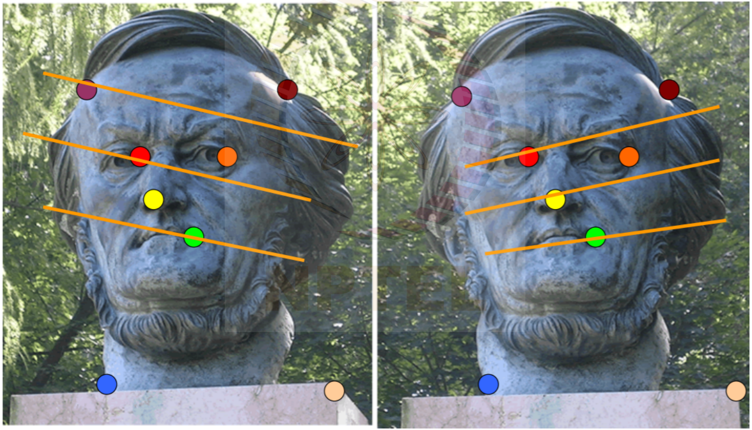
Rotation



Credit: Aaron Bobick, Washington University in St. Louis

RANSAC Applications

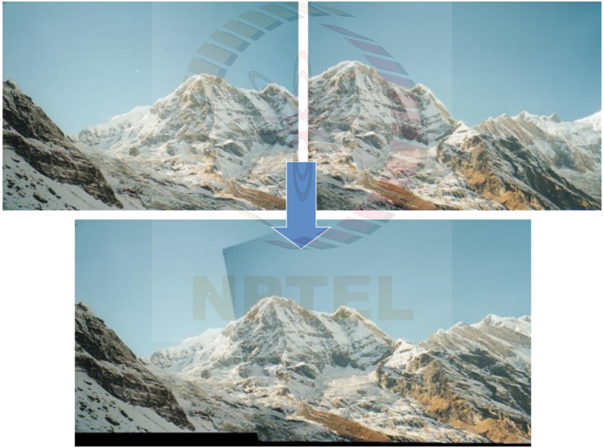
Estimating transformation matrix (also called **fundamental matrix**) relating two views



Credit: Derek Hoeim, UIUC

RANSAC Applications

Computing a **homography** (e.g., image stitching)



Credit: Ali Farhadi, Univ of Washington

Homework

Readings

- Chapter 4.3, 6.1, Szeliski, *Computer Vision: Algorithms and Applications*
- Papers on the respective slides (for more information)

References

- Martin A. Fischler and Robert C. Bolles. "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography". In: *Commun. ACM* 24.6 (June 1981), 381–395.
- Bruce D. Lucas and Takeo Kanade. "An Iterative Image Registration Technique with an Application to Stereo Vision". In: *Proceedings of the 7th International Joint Conference on Artificial Intelligence - Volume 2*. IJCAI'81. Vancouver, BC, Canada: Morgan Kaufmann Publishers Inc., 1981, 674–679.
- Richard Szeliski. *Computer Vision: Algorithms and Applications*. Texts in Computer Science. London: Springer-Verlag, 2011.
- *Avrithis, Yannis, Deep Learning for Vision (2018)*. URL: https://sif-dlv.github.io/ (visited on 05/21/2020).
- *Hoiem, Derek, CS 543 - Computer Vision (Spring 2011)*. URL: https://courses.engr.illinois.edu/cs543/sp2017/ (visited on 04/25/2020).