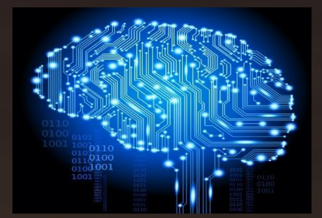


Trajectory (Motion) estimation of Autonomously Guided vehicle using Visual Odometry

Problem Statement

- Given sequence of images ,estimate trajectory of autonomous vehicle using visual odometry
- Aim is to find camera poses from set of images taken at discrete interval
- **Problem formulation:**

We have to find a Transormation matrix which relates two image frames i.e. how the two frames are rotated and translated from each other.

let set of images be $\{I_0, I_1, I_{2.....}I_{k-1,}I_k\}$,camera poses be $\{C_0, C_1, C_{2.....}C_{k-1,}C_k\}$

and transformation matrix is given by

$$T_{k,k-1} = \begin{bmatrix} R_{k,k-1} & t_{k,k-1} \\ 0 & 1 \end{bmatrix}$$

Where:

$T_{k,k-1}$ is homogenous transformation matrix between images I_k and I_{k-1}.

$R_{k,k-1}$, $t_{k,k-1}$ are rotation and translation matrix between images I_k and I_{k-1}.

Proposed Algorithm

Feature detection (SIFT/SURF/FAST)

↓

Feature Matching

↓

Outlier Removal using RANSAC

↓

Estimate motion using Essential Matrix

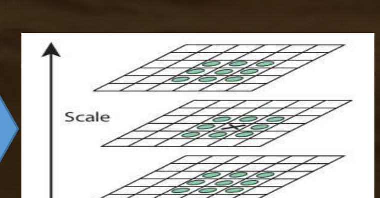
↓

Windowed bundle Adjustment (optional) X

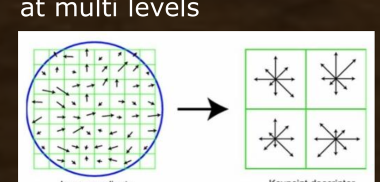
1. Feature Detection (SIFT)

•Gaussian of Images in 4 Octaves (Scale Space)

•Difference of Gaussians



•Local Maxima in neighborhood at multi levels



•Feature Descriptors(Extraction)

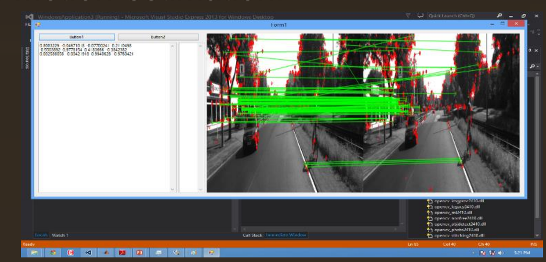
Save Features

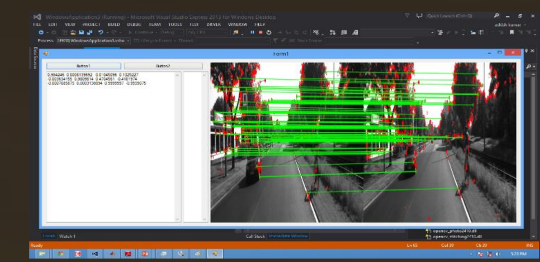
2. Feature Matching

After having feature descriptors from two images, use nearest neighbor approach. The point which is at minimum euclidean distance and distance is below a threshold is selected as correspondent of point in image 1. let point correspondences be x , x'

3. RANSAC Algorithm (Regression between Inliers)

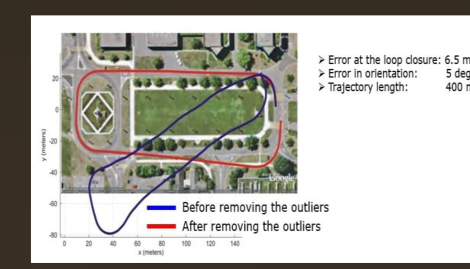
Usually these matched points are contaminated with "Outliers" i.e. wrong correspondences. Which can lead to errors in Motion estimation.





RANSAC is a non deterministic method. No of Inliers in the output depends on no of loops of sampling. As time increases Probability of convergence increases. No of loops Required are given by

$$N = \frac{\log(1-p)}{\log(1-(1-\varepsilon)^s)}$$



p= probability of success

ε = % of outliers in data

s = no of points by which model can be instantiated.

4. Motion Estimation:

To compute T matrix which is composed of Rotation matrix (R) and translation matrix (t), we need to compute Essential matrix using the Inliers filtered by RANSAC. R is 3x3 and t is 3x1 matrices.

E- matrix can be computed using RANSAC, Normalized 8 point algorithms. Let SVD of $E = UDV^T$ Then there are four solution. Let W be a skew symmetric matrix. Then four solutions are

$$W^T = \begin{bmatrix} 0 & \pm 1 & 0 \\ \mp 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad R = U(\pm W^T)V^T$$

$$t = \pm[u_{13}, u_{23,} u_{33}]$$

Only one of them corresponds to true configuration, which is obtained by 3D triangulation of image correspondences (Inliers). Camera pose is given by P=[R|t] . Epipolar constraint on E is given by

$$x'^T E x = 0$$

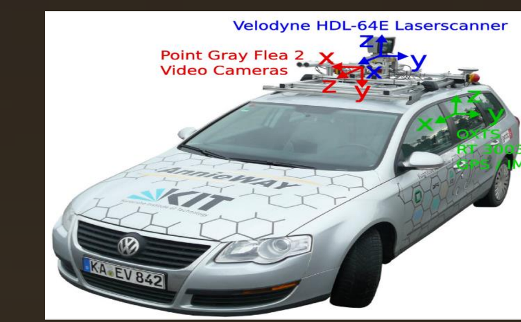
Triangulation

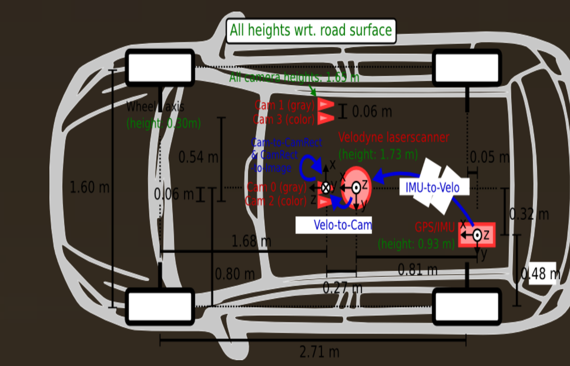
It is done by direct linear transform (DLT). Let P and P' be camera matrices corresponding to two images and p^{iT} and p'^{iT} be rows of P and P' matrices, where i = 1,2,3. Then matrix A is given by constraint x=PX. Here X is homogenous world coordinate.

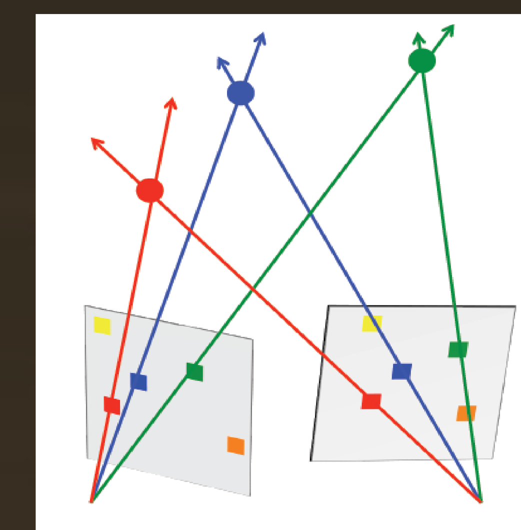
If SVD of A is UDV' then the 3D point is last column of V

Now , we choose correct configuration by imposing the constraint that 3D point must lie in front of both cameras.

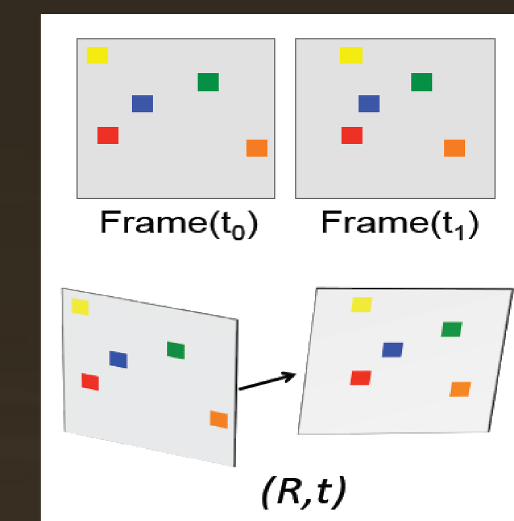
$$A = \begin{bmatrix} xp^{3T} - p^{1T} \\ yp^{3T} - p^{2T} \\ x'p'^{3T} - p'^{1T} \\ y'p'^{3T} - p'^{2T} \end{bmatrix} \qquad E = \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix} \begin{bmatrix} & R_{k,k-1} & \end{bmatrix}$$



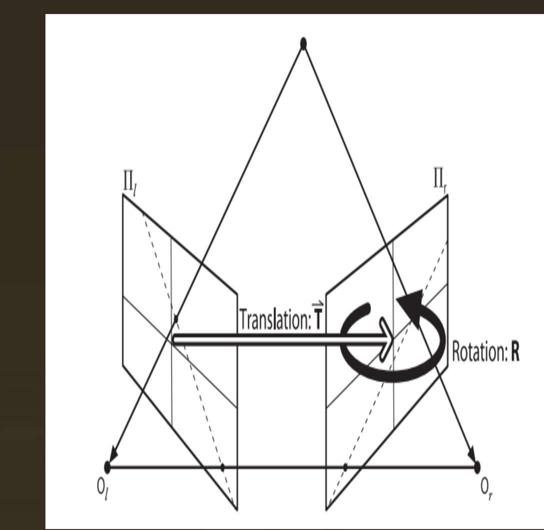




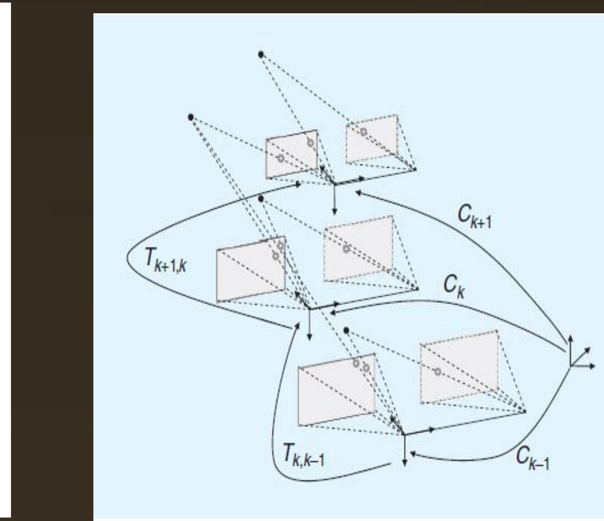
• 3D point correspondence



• Relation between to frames



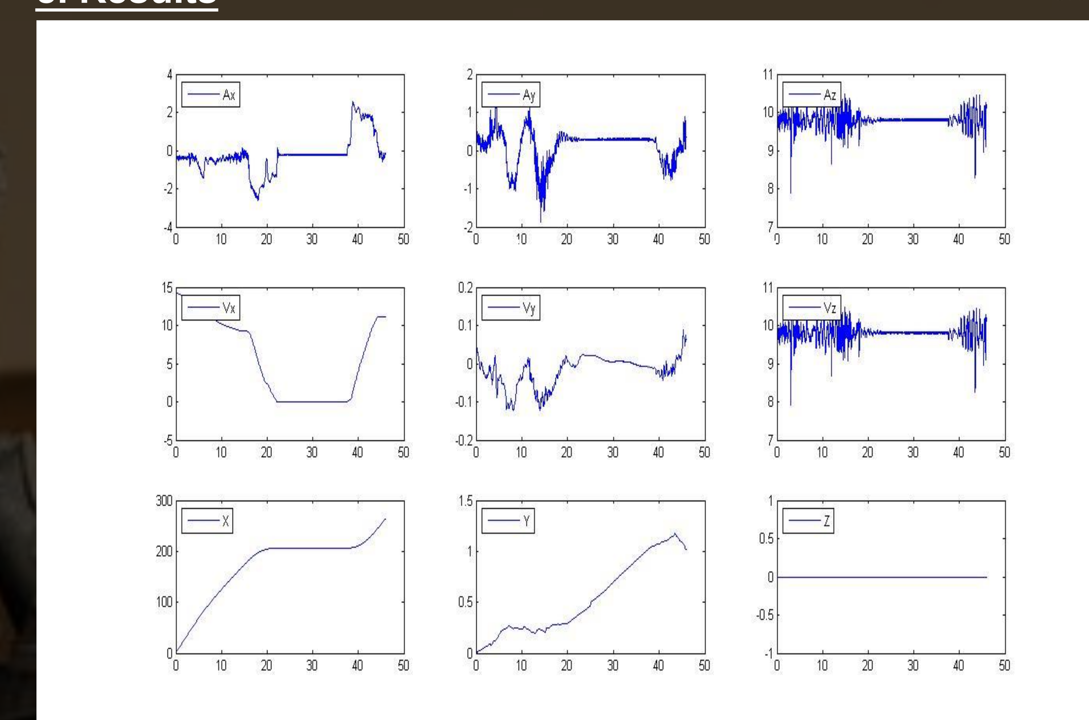
• Relation between to frames

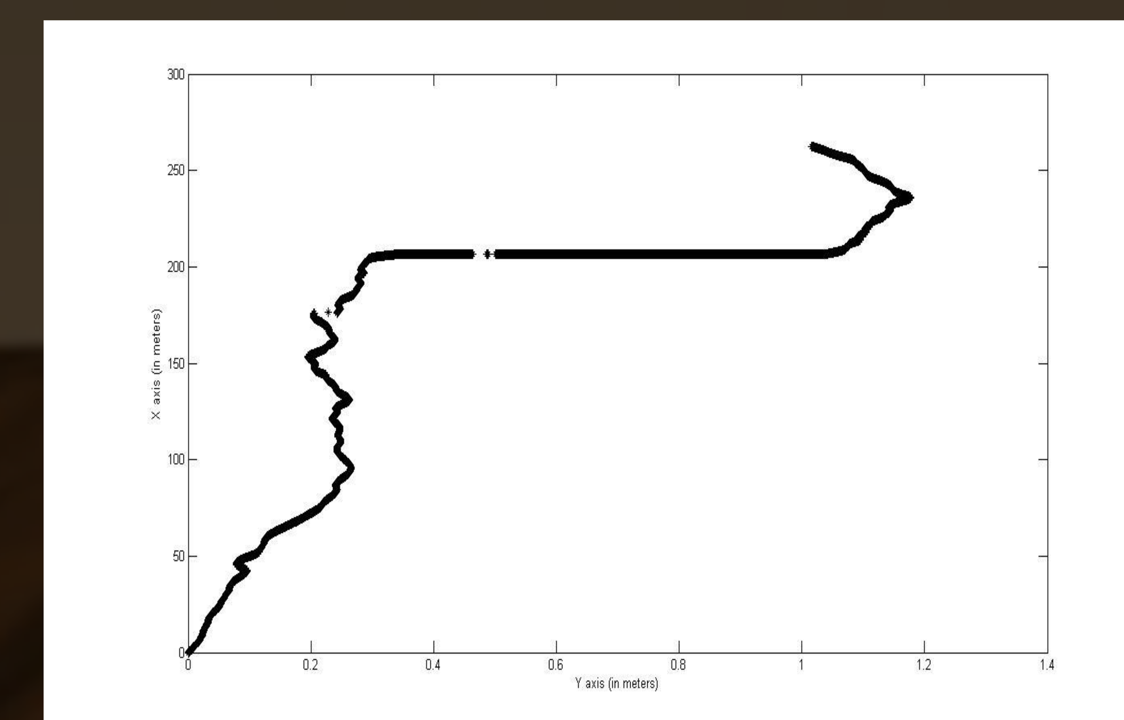


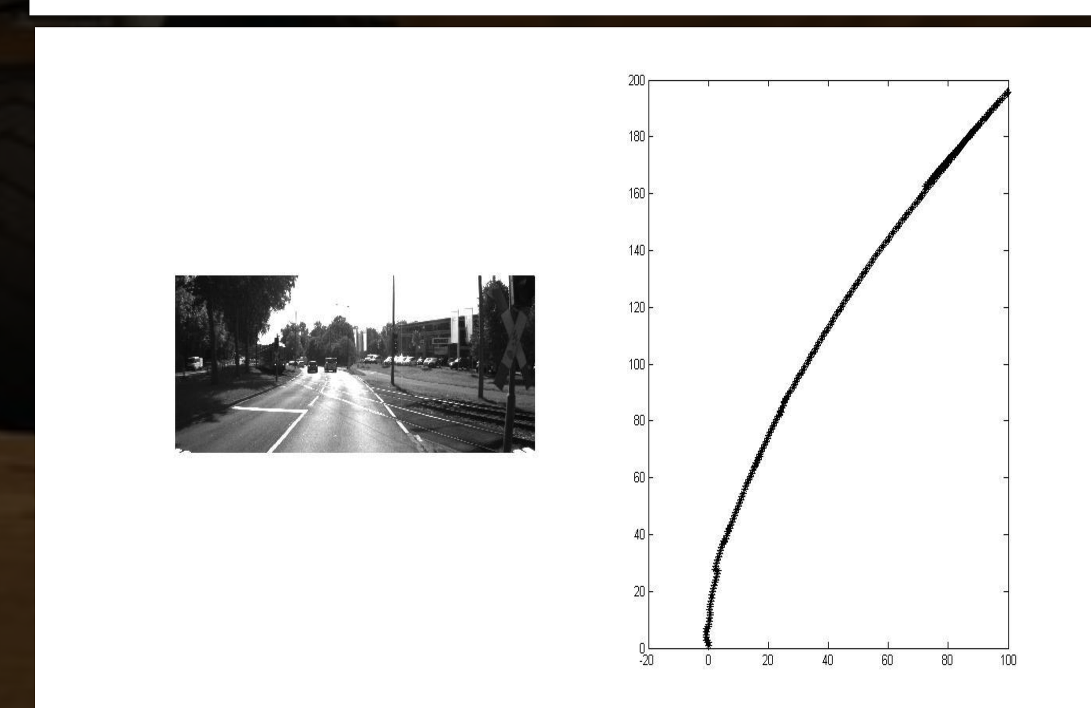
• Relation between to Camera Poses

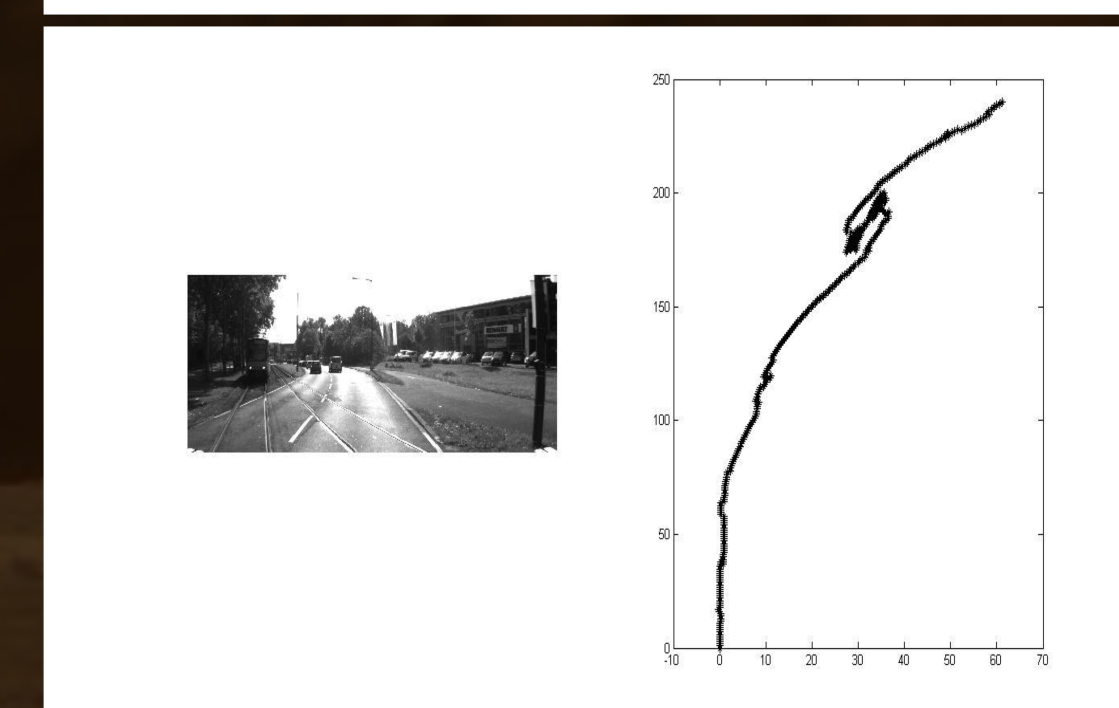
5. Results

Ground truth









Results of program written in Visual Basic with EmguCV

Results of program written in MATLAB

References:

[1]. Andreas Geiger, Philip Lenz, Christoph Stiller and Raquel Urtasun. Vsion meets Robotics: *The KITTI dataset.* In Journal "International Jouranl of Robotics Research" (IJRR); 2013

[2]. Scaramuzza, D., Fraundorfer, F., *Visual Odometry: Part I - The First 30 Years and Fundamentals*, IEEE Robotics and Automation Magazine, Volume 18, issue 4, 2011.

[3]. Fraundorfer, F., Scaramuzza, D., *Visual Odometry: Part II - Matching, Robustness, and Applications*, IEEE Robotics and Automation Magazine, Volume 19, issue 1, 2012.

[4]. David Niste´ r, Member, IEEE , *"An Efficient Solution to the Five-Point Relative Pose Problem"* ,IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 26, NO. 6, JUNE 2004

[5]. *Multiple View Geometry in Computer Vision 2nd Edition by* Richard Hartley Australian National University, Canberra, Australia and Andrew Zisserman University of Oxford, UK

[6]. H.C. Longuet, Higgins *"A computer algorithm for reconstructing a scene from two projections".*

Presented By- Ashish Kumar

Roll No. 14104023

M.Tech. 2nd Sem EE (2014-16)

CS365A (Artificial Intelligence)

Instructor- Prof. Amitabh Mukherjee