Write-up :- 3

1. **Single view camera calibration** : Camera calibration is the process of determining specific camera parameters in order to complete operations with specified performance measurements. It can be defined as the technique of estimating the characteristics of a camera.

**2. Stereo camera system**

A **stereo camera** is a type of [camera](https://en.wikipedia.org/wiki/Camera) with two or more lenses with a separate [image sensor](https://en.wikipedia.org/wiki/Image_sensor) or film frame for each lens. This allows the camera to simulate human [binocular vision](https://en.wikipedia.org/wiki/Binocular_vision), and therefore gives it the ability to capture three-dimensional images, a process known as [stereo photography](https://en.wikipedia.org/wiki/Stereo_photography).

Stereo cameras may be used for making [stereoviews](https://en.wikipedia.org/wiki/Stereoscopy" \o "Stereoscopy) and 3D pictures for movies, or for [range imaging](https://en.wikipedia.org/wiki/Range_imaging#Stereo_triangulation). The distance between the lenses in a typical stereo camera (the intra-axial distance) is about the distance between one's eyes (known as the intra-ocular distance) and is about 6.35 cm, though a longer base line (greater inter-camera distance) produces more extreme 3-dimensionality.

3D pictures following the theory behind stereo cameras can also be made more inexpensively by taking two pictures with the same camera, but moving the camera a few inches either left or right. If the image is edited so that each eye sees a different image, then the image will appear to be 3D. This method has problems with objects moving in the different views, though works well with still life.

Stereo cameras are sometimes mounted in cars to detect the lane's width and the proximity of an object on the road.

Not all two-lens cameras are used for taking stereoscopic photos. A [twin-lens reflex camera](https://en.wikipedia.org/wiki/Twin-lens_reflex_camera) uses one lens to image to a focusing/composition screen and the other to capture the image on film. These are usually in a vertical configuration.

**3. Two-view geometry/ Epipolar Geometry**

**Point correspondence :**

The Particle Based Modeling (PBM)• approach to SSM constructs a correspondence-point model of shape, which describes shape variation by choosing a discrete set of corresponding points on shape surfaces whose relative positions can be statistically analyzed. Each of the points is called a correspondence point.

**Homography** :

Homography, also referred to as planar homography,• is a transformation that is occurring between two planes. In other words, it is a mapping between two planar projections of an image. It is represented by a 3x3 transformation matrix in a homogenous coordinates space.

**Fundamental matrix :**

The fundamental matrix is a 3×3 matrix• which relates corresponding points in stereo images. In epipolar geometry, with homogeneous image coordinates, x and x′, of corresponding points in a stereo image pair.

Fx describes a line (an epipolar line) on which the corresponding point x′ on the other image must lie.

That means, for all pairs of corresponding points holds

x’TFx = 0

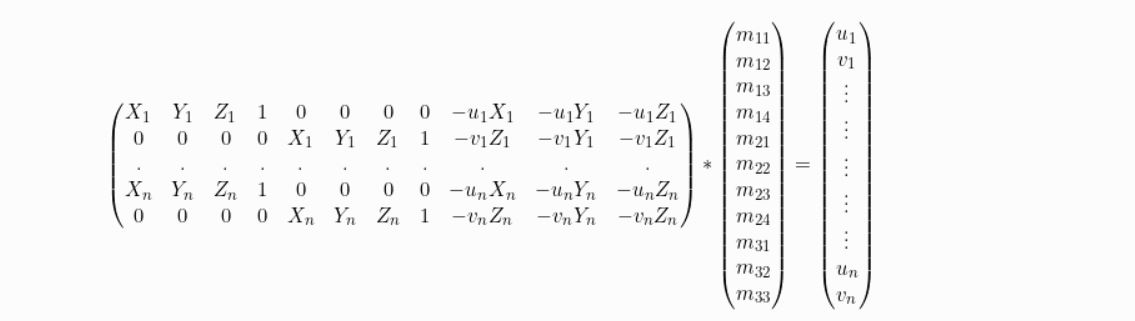
Being of rank two and determined only up to scale, the fundamental matrix can be estimated given at least seven point correspondences. Its seven parameters represent the only geometric information about cameras that can be obtained through point correspondences alone.

**Estimating the fundamental matrix using normalised 8-• point algorithm :**

The normalized eight-point algorithm is used to compute the fundamental matrix given point correspondences x = (u, v) and x' = (u', v') in two images. Each point correspondence generates one constraint on the fundamental matrix F and must satisfy the epipolar constraint equation.

https://latex.codecogs.com/gif.latex?\inline&space;\bold%7bx%27%7d%5eTF\bold%7bx%7d&space;=&space;0.

Expanding the matrices out by multiplication, we obtain the following equation for n point correspondences



where A is the n x 9 equation matrix, and f is a 9-element column vector containing the entries of the fundamental matrix F. From here, the least-squares solution f is easily computed by performing singular value decomposition (SVD) on the matrix A=UDVT. It is well-known that the vector f that minimizes ||Af|| such that ||f|| = 1 can be found along the column of V corresponding to the least singular value. Next, we rearrange the 9 entries of f to create the 3x3 fundamental matrix F. Then, we perform SVD on F to obtain

https://latex.codecogs.com/gif.latex?F&space;=&space;U_fD_f%7bV_f%7d%5eT

**Computing camera matrices from fundamental matrix :**

camera matrix or projection matrix can be computed as :

P = K \* [R | t] or P = [ [e']x \* F | e']

If we computed F ( fundamental matrix ) from the 8 points algorithm, then you can recover only projective geometry up to a 3D homography (i.e. a 4x4 transformation).

To upgrade to euclidian space, there are 2 possibilities, both starting by computing the essential matrix :- 1. First possibility is to compute the essential matrix from F: E = transpose(K2)\*F\*K1.

2. Second possibility, is to estimate directly the essential matrix for these 2 views: Normalize 2D points by pre multiplying with inverse of K for each image ("normalized image coordinates"). Apply the (same than for F) 8 points algorithm on these normalized points. Enforce the fact that the essential matrix has its 2 singular values equal to 1 and last is 0, by SVD decomposition and forcing the diagonal values. Once we have the essential matrix, we can compute the projection matrix in the form

P = K \* [R | t]

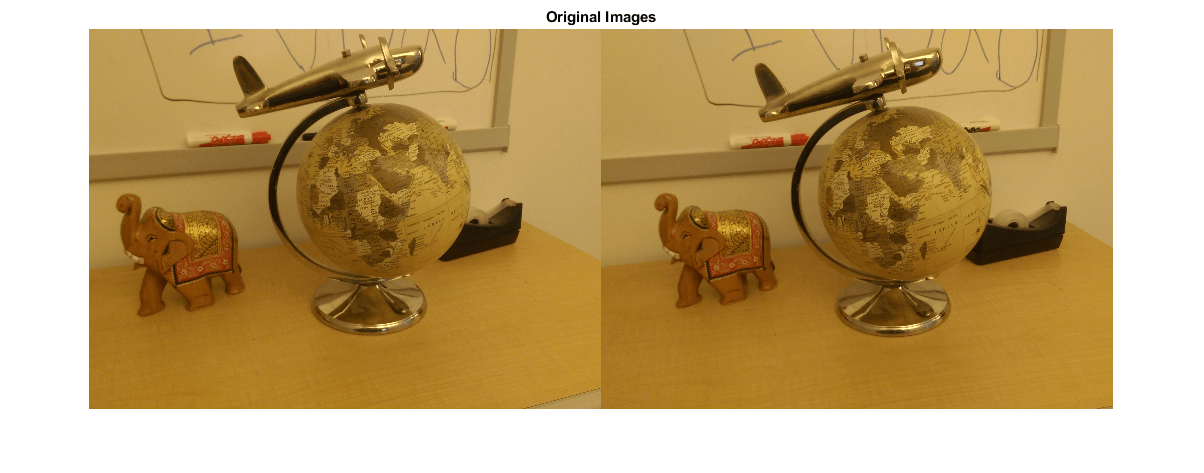
R and t can be found thanks to the elements of the SVD of E . However, we will have 4 possibilities. Only one of them projects points in front of both cameras, so we shall test a point (if you are sure of it) to remove the ambiguity among the 4. And in this case we will be able to place the camera and its orientation (with R and t of the projection) in our 3D scene.

**4. Two-view Structure from Motion / Two-view Reconstruction**

Structure from motion (SfM) is the process of estimating the 3-D structure of a scene from a set of 2-D images. This example shows you how to estimate the poses of a calibrated camera from two images, reconstruct the 3-D structure of the scene up to an unknown scale factor, and then recover the actual scale factor by detecting an object of a known size.

This example shows how to reconstruct a 3-D scene from a pair of 2-D images taken with a camera calibrated using the [Camera Calibrator](https://www.mathworks.com/help/vision/ref/cameracalibrator-app.html) app. The algorithm consists of the following steps:

1. Match a sparse set of points between the two images. There are multiple ways of finding point correspondences between two images. This example detects corners in the first image using the detectMinEigenFeatures function, and tracks them into the second image using vision.PointTracker. Alternatively you can use extractFeatures followed by matchFeatures.
2. Estimate the fundamental matrix using estimateEssentialMatrix.
3. Compute the motion of the camera using the relativeCameraPose function.
4. Match a dense set of points between the two images. Re-detect the point using detectMinEigenFeatures with a reduced 'MinQuality' to get more points. Then track the dense points into the second image using vision.PointTracker.
5. Determine the 3-D locations of the matched points using triangulate.
6. Detect an object of a known size. In this scene there is a globe, whose radius is known to be 10cm. Use pcfitsphere to find the globe in the point cloud.
7. Recover the actual scale, resulting in a metric reconstruction.



This example showed how to recover camera motion and reconstruct the 3-D structure of a scene from two images taken with a calibrated camera.