Summary of Algorithms

Five Point Algorithm

In this section, I will dissect what is going on inside the Five-point Algorithm that was provided as it is for this assignment. The five-point algorithm by Nister has a few steps.

First of all the preliminaries should be known. So let us take a camera with some know parameter K. Then we know that two corresponding images in a stereo pair can be compared with each other using the fundamental matrix F.

$$\mathbf{x^T\ F\ x = 0}$$

The fundamental matrix can be changed to Essential matrix E if we have calibration information. Such that

$$\mathbf{K^{-T}\ E\ K^{-1} = F} \quad \textbf{(equation 1)}$$

This matrix has five degrees of freedom. So it will satisfy the the following also:

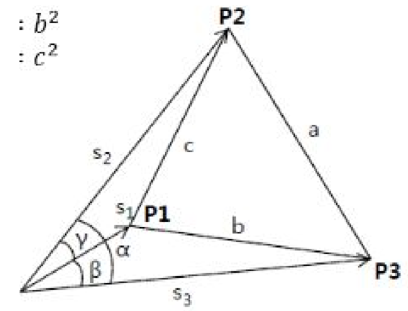
$$\mathbf{2EE^TE-tr(EE^T)E = 0} \quad \textbf{(equation 2)}$$

The combination of the above equation(2) makes nine equations and 5 equations from the five points correspondences can be obtained plus *det(F) = 0* is also an equation, means we have all the equations we need to solve the Essential matrix.

Now with the above equation known the Nister algorithm First gets the null space representation of all the five correspondence points. You can do that by using equation $x^T\ F\ x = 0$. Then we need to use the nine equation formed by equation 2. We can use these equations to form a 9 x 20 coefficient matrix. The unknowns in the equation are [x^3 ,y^3 ,x^2 y^1 ,x^1 y^2 ,x^2 z^1 ,y^2 z,y^2 xyz,xy,xz,x,yz,y,z,1,x^2 z,y^1 z^2 ,z^3 z^3]. Next Gauss Jordan Elimination can be applied to the 9 x 20 matrix to make it upper triangular. Now we can use the *det(F)* equation and extract the determinants of the 4 x 4 matrices and then again perform Gauss Jordan Elimination. This gives us a 10th-degree univariate polynomial which can be solved to obtain 10 solutions for z dimensions. Now we use back substitution to solve for the other unknowns. This will give us the essential matrix.

Three-Point Algorithm P3P

To understand the p3p algorithm lets us first understand what is the problem. Imagine we have a camera then we can establish vectors between image points and world points. Once we know these vector we can also find the angles between the image point and world points. This can be visualized as follows in Figure 1.



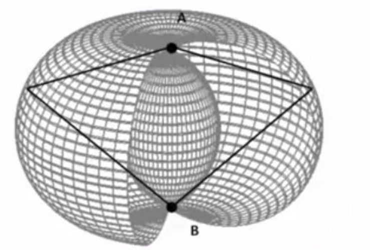


Figure1 Figure2

Once we know these vector and the angles between them. The question is how to find the camera pose?

The p3p says we need at least three points to find the pose. In 2D or 3D configuration. If we use only two points and we use only the angles, we can not constrain the position of the camera. In the case of 2D, using 2 points the camera can lie anywhere on a fixed circle. While in case of 3D the camera can lie anywhere on the surface which is called toroid as in Figure 2. So two points are not enough in 2D or 3D. By adding a third point, we can constrain the pose of the camera just by using the angles defined. Even though the solution will not be unique but there will be a finite number of solutions. This is called the **resection problem in Photogrammetry**. This is a very old problem and the first researcher who worked on this were Snellius and Pothymous. The full solution was first given by **Gunners** in **1841** in **Germany**. In computer vision, it corresponds exactly to the problem where we have three projections of points and if we know the calibration, we can find the angles between the arrays.

In Figure 1 we have four triangles which we should consider. The first one is the triangle formed between the three world points with side a,b and c. The other three triangles are the ones which are formed between two world points and the images points. It can be clearly seen by inspection that there will be three such unique triangles. Then since we know all the world points and images points and the angles. We can apply the law of cosines for all three triangles between two world points and the image points as follows.

$$s_2^2 + s_3^2 - 2s_2 s_3 \cos\alpha = a^2$$

$$s_1^2 + s_3^2 - 2s_1 s_3 \cos\beta = b^2$$
$$s_1^2 + s_2^2 - 2s_1 s_2 \cos\gamma = c^2$$

In the above equations we just need to find the distance s1,s2,s3 and after that, we can easily find the Projection matrix. In the above equation we make the following substitution :

$$s_2 = us_1 \text{ and } s_3 = vs_1.$$

Then we can get the following three equations where we have s1 on the left side :

$$\begin{aligned} s_1^2 &= \frac{a^2}{u^2 + v^2 - 2uv\cos\alpha} \\ &= \frac{b^2}{1 + v^2 - 2v\cos\beta} \\ &= \frac{c^2}{1 + u^2 - 2u\cos\gamma} \end{aligned}$$

Combining these three we can form two equation in terms of u and v as follows:

$$\begin{aligned} &u^2 + \frac{b^2 - a^2}{b^2}v^2 - 2uv\cos\alpha \\ &\qquad + \frac{2a^2}{b^2}v\cos\beta - \frac{a^2}{b^2} = 0 \\ &u^2 - \frac{c^2}{b^2}v^2 + 2v\frac{c^2}{b^2}\cos\beta \\ &\qquad - 2u\cos\gamma + \frac{b^2 - c^2}{b^2} = 0. \end{aligned}$$

In the above equation, we know all the angles and the length of triangles a, b and c. We only need to solve this to find u and v. After which we can find s1,s2 and s3.
After finding s1,s2 and s3, we can find the **absolute orientation ,** we need to find rotation and translation matrix in the following equation

$$p_i = Rp_i' + T \qquad i = 1,2,3$$

In the above equation, we already found the *pi* which is s1, s2 and s3. Simply by solving the above equation as AX= B, in a linear fashion, we can find the pose and translation matrix to determine the complete projection matrix P.